View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Padua@thesis

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di Laurea Magistrale in Fisica

Tesi di Laurea

Dynamics of Quantum Spin Chains

Relatore

Prof. Antonio Ponno

Laureanda

Alessandra Colla

Anno Accademico 2017/2018

Abstract

In the present thesis we perform a step towards understanding the effectiveness of Hamiltonian perturbation theory in quantum dynamics, which is made possible by the Hamiltonian framing of quantum mechanics. To such a purpose, we focus our attention on quantum spin chains, namely on the isotropic Heisenberg ferromagnet, which is well known to be equivalent, via the Jordan-Wigner transformation, to a system of interacting fermions, and is also known to be exactly integrable through the Bethe ansatz. The interaction term of the Hamiltonian, of order four in fermionic creation and annihilation operators, is treated as a perturbation in the regime of small excitations: small number of reversed spins or, equivalently, small number of fermions per site. The calculation of the first order normal form Hamiltonian amounts to erasing all the nonresonant interaction terms, and allows to draw interesting conclusions on the dynamics of the system, e.g. to determine approximately conserved quantities. Moreover, we get a good correspondence between our perturbative energy spectrum and the exact one computed via the Bethe ansatz, the approximation being of course the more accurate the smaller is the number of fermions per site. The present method, here tested in an exactly solvable case, can be applied to any nonintegrable system of weakly interacting fermions (and/or bosons).

Contents

Introduction

One of the many advantages that came with the advent of quantum mechanics was finding the explanation to ferromagnetic phenomena, a true mystery at the time. In fact, purely classical physics fails to explain ferromagnetism from an atomic point of view, since it misses a proper concept for spin. The revolution came from Heisenberg, who found that the laws of quantum mechanics imply the existence of an effective interaction between electron spins on neighboring atoms with overlapping orbital wave functions. This exchange interaction, usually of the form

$$g_{ij}\,\hat{\vec{S}}_i \cdot \hat{\vec{S}}_j \quad ,$$

is caused by the combined effect of the Coulomb repulsion and the Pauli exclusion principle. Starting from this discovery, many other spin models arose to comprise a microscopic theory of ferromagnetism.

Spin chains

Spin chains are one-dimensional spin models. Because it is easier to deal with complex structures in lower dimensions, they are the most studied class of spin models. Their application is not restricted to just ferromagnetic models: they have a vast range of application due to their core structure and versatility. In modern applications, they are of crucial importance in fields like quantum information theory and computability theory.

Classical spin chains

The classical versions of spin chains were introduced as semi-classical descriptions of magnetic phenomena in metals, particularly one-dimensional ones. The reason we say it is a semi-classical approach is that there is actually no real concept of spin in classical physics, being it a purely quantum feature. That said, it was possible to study the phenomena resulting from spin presence using classical tools, by modeling the spin as a vector that is intrinsic

to each particle and of fixed norm. Depending on the kind of interaction between the spins (for example, nearest neighbors), but also on the vector space to be spanned by the vectors, we have different models with different characteristics. To mention a few: the Ising model for vectors pointing in just one direction, the XY model for planar interaction and fixed z component, the Heisenberg model for vectors pointing on a sphere, with interaction along each direction.

For big values of spin S, the classical description is more than satisfactory; indeed, as S goes to infinity, the quantum models actually revert to the classical ones. From the classical approach, it is also possible to obtain the *Landau-Lifshitz equation* — and we will do just so in the first chapter —, which can be then modified to become the more precise *Landau-Lifshitz-Gilbert equation* (LLG) [1]. Together with the Maxwell equations, this equation is the one describing the dynamics of basically any magnetic device. There are some other equations that are useful in magnetism like the Bloch equation, the Landau-Lifshitz-Bloch equation, or the Ishimori equation, all of them in some way deriving from the Landau-Lifshitz. In fact, the explicit form of the Gilbert equation is similar to the original LL equation, therefore the name. Both equations describe the dynamics of a single spin in a lattice, with the LLG equation being the one describing the correct physics while the LL equation fails in the limit of consistent damping [2].

Quantum spin chains

Quantum spin-chains are the quantum (therefore more accurate) version of the classical spin chains just mentioned. In this case, the spin is always considered having three components, represented by the Pauli matrices, as it should be, and the type of model considered is identified by the kind of interaction present — Ising for interaction along z only, XY model for interaction along x and y, Heisenberg for interaction on all components, all of these being nearest neighbor interactions. Naturally, the quantum formulation also changes how we look at the states of the chain, meaning that we allow complex linear superpositions of different spin configurations as a physical state.

Some quantum spin chains are exactly solvable, i.e. they are "quantum integrable" systems in 1+1 spacetime dimensions. The dynamics of the system, i.e. how a particular state evolves in time, is governed by the Schrödinger equation, if we work under the Schrödinger picture. In this dissertation, we will work in the Heisenberg picture frame, since it is more suitable for constructing a Poisson algebra on the system, and we will therefore look at the evolution of the operators involved in order to determine the dynamics of the

system.

"Exactly solvable" models or "quantum integrable" systems are important because they may be non-trivial interacting systems where exact solutions can be obtained (for all values of the coupling constant). In this project, we will focus our attention on the quantum isotropic Heisenberg chain, or XXX chain, with Hamiltonian

$$\hat{H} = -g \sum_n \hat{\vec{S}}_n \cdot \hat{\vec{S}}_{n+1} \quad ,$$

which is an interactive integrable system. Various aspects of quantum spin chains, many of which are here useful, have been collected in [3].

Integrable systems

The concept of integrability has its roots in the context of differential equations, and bears the meaning of "to solve from initial conditions"; i.e., a system of differential equation is integrable if it is possible to integrate its equations using the initial conditions, obtaining the behavior of its variables. Integrability is a feature closely related to the presence of many conserved quantities, or first integrals, in the system. However, there are some other things besides this that can suggest integrability, for example the presence of algebraic geometry in the system, which can open paths to algebraic integrable techniques. Then, of course, a system is integrable whenever we are able to find an explicit solution. There are many interesting systems in physics that are integrable, like models for shallow water waves, described by the Korteweg-de Vries equation, the nonlinear Schrödinger equation, and the Toda lattice in statistical mechanics.

When we deal with classical Hamiltonian systems, the Liouville-Arnold theorem is a powerful tool to approach the concept of integrability, [4]. Unfortunately, in the quantum case there is no clear analogue of such a theorem, and it is much harder to understand if we are dealing with an integrable system or not. In the quantum setting, functions on phase space are to be replaced by self-adjoint operators on a Hilbert space, while the Poisson bracket is replaced by the Heisenberg commutator (i.e. the commutator over $i\hbar$). If we consider a system of particles, we might say that a quantum system is "integrable" if the dynamics are two-body reducible. This is because being a two-body reducible system equals to fulfilling the Yang-Baxter equation [5], the latter leading to trace identities that provide an infinite set of conserved quantities. This approach to quantum integrability is generally applied in the so called *quantum inverse scattering method*, introduced by Faddeev in 1979

(see [6]), together with the algebraic Bethe ansatz, to obtain exact solutions in many quantum spin chains. Apparently, this method can be applied in one-dimentional quantum systems only. Another way to study integrability of a quantum system is to see what happens when it is possible to map it into a system of particles — from here, the name *fermionization* or *bosonization*, depending on the type of particle we map the system into. Clearly, if the system is mapped into free particles, then it is one-body reducible, i.e. naturally integrable. This happens for example — and we will see this later in detail — for the XY model.

Examples of quantum integrable models are the Lieb-Liniger Model, the Hubbard model [7] and several variations on the Heisenberg model. For more on quantum integrability, see [8], [9], [5].

Just as the non-integrability in classical systems is related to deterministic chaos theory, non-integrability in quantum systems is related to the somewhat less precise concept of quantum chaos. For a quick study on these concepts, see [10].

Integrable techniques

As we just mentioned, there is more than one way to tackle the problem of quantum integrability. In this dissertation, we will not delve into the quantum inverse scattering method; we will, however, give a review of the coordinate Bethe ansatz technique (Bethe's original work) and perform our calculations in the setting of fermionization.

Bethe ansatz

Historically, Bethe's work [11] on the isotropic case ($g_x = g_y = g_z$) of the Heisenberg model, known as the XXX chain, had unexpected success and was the starting point in the development of more modern techniques for solving quantum systems. The "ansatz" consists in thinking the eigenstates of the XXX spin-chain as a superposition of plane waves. The momenta describing these plane waves end up satisfying a set of non-linear equations, called Bethe's equations, which are usually solvable via computational methods. In literature, this approach is referred to as "coordinate Bethe ansatz", to separate it from its algebraic counterpart, and its core idea has been applied to solve many other quantum integrable systems. We give a more detailed account of the original coordinate Bethe ansatz in Appendix B. For more on it, see [11], [12], [13], [14].

Fermionization or bosonization

This kind of approach is, of course, particularly interesting, since it means that anything we understand about the particle version of the system can be immediately applied to learn something about the spin system. The idea is that particle systems are easier to deal with, therefore the best way of understanding spin systems may be to map them onto mathematically equivalent but physically different systems, whose properties we already understand. In other words, we introduce a quasiparticle description of the spin system, in order to simplify its understanding. This idea became central in various aspects of modern physics, including the field of superconductivity and the quantum Hall effect.

Whether we map into bosons or fermions usually depends on the properties of the system, and what they are most similar to. For example, the spin 1/2 Heisenberg chain is naturally mapped into a system of interacting fermions via the Jordan-Wigner transformation [15]. A case of bosonization is instead the Luttinger model, which can be mapped into a system of non-interacting bosons [16]; its fermionic version (interacting) is treated perturbatively in [17]. We therefore see that this method is very versatile and that there are multiple paths that one can take.

Another reason why this mechanism is so important is that it is possible to generalize it to more than one dimension, so that we can also study spin lattices. To give a few examples, Kitaev in [18] provides a fermionization for the 2d Luttinger model, while recently Chen, Kapustin and Radičević described in [19] a 2d analog of the Jordan-Wigner transformation while also giving some examples of 2d bosonization. This is actually nothing new, as Lieb, Schultz and Mattis presented, already in 1961, an alternative to Onsager's 2d Ising model solution using the fermionization technique [20].

In this discussion, we will use this method by fermionizing the Heisenberg isotropic chain via the Jordan-Wigner transformation obtaining a Hubbard-like model [7], and we will exploit the fact that the XY part of the Hamiltonian is mapped into a free fermion system, and is therefore integrable, to apply perturbation theory; the free system will constitute the unperturbed Hamiltonian, while the remaining term will be treated as a perturbation.

Perturbation theory

A powerful tool that constitutes the central action in this dissertation is a particular version of Hamiltonian perturbation theory, which will be taken for granted in the main discussion but explained in detail in Appendix A.

Expanding the classical and well known version of Hamiltonian perturbation theory, the approach we use has it roots in the Poisson algebra framework of Hamiltonian systems. The core idea of this framework is that of extending the definition of a Hamiltonian system to a more general class of systems by admitting other forms of Poisson structure besides the usual symplectic one when studying the dynamics of the system. The theory for this, together with the properties the tensor must satisfy in order to be suitable for Hamiltonian formulation, can be found in [21], [22].

Once we adopt the Poisson algebra formalism, we can exploit the algebraic structure of quantum mechanics given by the replacement of the Poisson bracket with the Heisenberg commutator to effortlessly transfer the mechanism and results of classical perturbation theory directly onto the quantum setting, and to therefore use it upon quantum systems.

The main goal of this kind of perturbation theory is that of finding a suitable unitary transformation which enables us to eliminate (partially or completely) the perturbation. The unitary transformation will actually be a composition of Schrödinger time evolution operators generated by unknown Hamiltonians, which we call *generating Hamiltonians*. In particular, this method is a formal and justified way to eliminate the less important terms in the perturbation, usually the same contributions that physicists neglect because they intuitively know them to be of less impact. The central result of our Hamiltonian perturbation theory is what we call *Averaging principle*, namely a theorem demonstrating the possibility to rewrite the perturbation in "normal form" up to a pre-fixed order, where each term considered has the remarkable property of being a first integral for the unperturbed system.

Our use of Hamiltonian perturbation theory will be exactly to profit from its applicability in quantum mechanics by exploiting it on the fermionized Heisenberg isotropic spin chain, where the role of the (easily integrable) unperturbed system will be played by the terms that are quadratic in fermionic creation and annihilation operators, while the quartic term will be treated as a perturbation. As we will see, this is allowed when the quartic term has an impact on the dynamics which is small with respect to the contribution from the unperturbed one; this happens when the number of reversed spins (as in, reversed in relation with the aligned ground state) is small when compared to the available number of sites.

Structure and results

Before we start the discussion, let us anticipate the way this thesis is organized, and the main results we obtained.

- The first chapter serves as an overlook on the classical Heisenberg model, along with the study of its dynamic. From the equations of motion, the continuum limit provides us with a version of the already mentioned Landau-Lifshitz equation.

- In the second chapter, we introduce our quantum model of the Heisenberg chain. Starting from the quantized version of the Hamiltonian, we go through the ladder operators $\hat{S}^{\pm}$ to then formulate the Hamiltonian in terms of fermionic creation and annihilation operators on the sites via the Jordan-Wigner transformation, completing the mapping from spin chain to a system of interacting spinless fermions. We dedicate a section to the concepts of the general technique by Jordan and Wigner.

- In the third chapter, the core of the discussion, we apply perturbation theory to the system, its applicability regimes being discussed in Section 3.4. First, we begin by changing coordinates via Fourier, finding spinless fermions labeled by momenta. In this setting, the unperturbed quadratic terms in the Hamiltonian are diagonalized, thus easily solved. In Section 3.2 we use the known dynamics of the unperturbed system to find the normal form of the Hamiltonian at first order, commenting on the results. While the dynamics in the case of odd number of sites is somehow trivial, the case of even number of sites is much more complicated, presenting additional interacting terms and a higher degree of degeneracy. In Section 3.3 we find the approximated values for the eigenvalues of our system with the help of quantum perturbation theory (eigenvalue corrections to first order), which is particularly necessary in the interacting case of even sites. Since the eigenvalue corrections for this case are not easily put in general form, we proceed in Section 3.5 by explicitly studying the corrections in the convenient yet not trivial example of $N = 6$, where N is the number of sites. In Section 3.6, we make an attempt at comparing the eigenvalues found via our perturbative approach with the exact ones found by the Bethe ansatz solution. The comparison reported is performed computationally for the case of two excitations, and for odd N, obtaining satisfactory results for bigger values of N, as expected. The other instances present more challenges, which we briefly discuss.

- Appendices A and B are dedicated to the illustration of the techniques of Hamiltonian perturbation theory and coordinate Bethe ansatz, respectively.

Chapter 1

The classical model

The classical formulation for 1D spin chains, but generally for any spin lattice, has served the purpose of modeling the physical phenomenon of magnetization and all that comes associated with the response to an external magnetic field. This is done by modeling microscopic magnetic moments (the spin) as vectors situated on each lattice site, which interact among themselves and with the mentioned external magnetic field. Such interaction is given by the scalar product between the quantities — indeed, naturally, the energy for a magnetic moment interacting with an external magnetic field is $-\vec{\mu} \cdot \vec{B}$. The exact form of the interaction can vary greatly, especially among the spins — long range, short range, negative or positive interaction etc. — the choice is made upon the physical properties of the material one wishes to study, be it ferromagnetic or anti-ferromagnetic, for example.

Assuming we have rescaled the system to dimensionless quantities, we call the spin variable placed at the nth site $\vec{S}_n$ and we assume it to have fixed magnitude (we are considering, in the quantum system part of the dissertation, a spin 1/2 chain). The external field will be denoted by $\vec{h}_n$, as it can virtually have different values on each site (it will not, in our case). Then, the total energy of the system, the starting point for the construction of a suitable Hamiltonian, will be:

$$H = -\frac{1}{2} \sum_{n,m} \vec{S}_n \cdot I_{nm} \vec{S}_m - \sum_n \vec{h}_n \cdot \vec{S}_n \quad . \tag{1.1}$$

Notice that the indexes n and m in the sums cover the whole lattice, which in our discussion will be made up of N sites with periodic conditions ($\vec{S}_{N+n} = \vec{S}_n \ , \ \forall n \in \mathbb{Z}$). The second term clearly describes the interaction of each spin with the external field, while the first represents the spin-spin interaction — meaning, $\sum_m I_{nm} \vec{S}_m$ is the total magnetic field that the spins generate

at site n, with I_{nm} being the symmetric interaction matrix. The factor $1/2$ in front of the first term is there simply to avoid recounting. Generally we require no self-interaction, i.e. we will have that $I_{nn} = 0$, while we will call the interaction *isotropic* if we have $I_{nm} = g_{nm}\mathbb{1}_3$, with g_{nm} a scalar matrix.

1.1 Poisson structure

In order to study the dynamic of the spin chain above described in the work environment of a Hamiltonian approach, and to exploit its various benefits, one has to define a Poisson structure for the variables that come in play. One can then treat the energy as a proper Hamiltonian, and thus be able to compute the equations of motion via the proper Poisson tensor. We will now do exactly this. For those who are not familiar with the concept, please see [22], [21].

1.1.1 General Heisenberg 1D spin chain

The appropriate Poisson structure for our spin system is found quickly once we realize that we can treat the spin just as a regular angular momentum. We can deduce it by analogy with the case of the Euler equations for the rigid body. In that case, the Poisson bracket for angular momenta is

$$\{L^i, L^j\} = -\sum_{k=1}^{3} \epsilon_{ijk} L^k \quad , \tag{1.2}$$

where the negative sign in front of the sum is purely conventional, as the map $\vec{L} \to -\vec{L}$ brings it to a plus sign.

Thus, the Poisson bracket for spin systems is simply

$$\{S_n^i, S_m^j\} = \delta_{nm} \sum_{k=1}^{3} \epsilon_{ijk} S_n^k \quad . \tag{1.3}$$

Naming J the Poisson tensor, the equations of motion follow accordingly:

$$\begin{aligned} \frac{\mathrm{d}S_n^i}{\mathrm{d}t} = (J\nabla H)_n^i = \sum_{m,j} \{S_n^i, S_m^j\} \frac{\partial H}{\partial S_m^j} = \sum_{m,j} \delta_{nm} \sum_{k=1}^{3} \epsilon_{ijk} S_n^k \frac{\partial H}{\partial S_m^j} = \\ = \sum_{kj} \epsilon_{ijk} S_n^k \frac{\partial H}{\partial S_n^j} = -\sum_{kj} \epsilon_{ikj} S_n^k \frac{\partial H}{\partial S_n^j} = -(\vec{S}_n \times \frac{\partial H}{\partial \vec{S}_n})^i \quad , \end{aligned}$$

encoded in

$$\frac{\mathrm{d}\vec{S}_n}{\mathrm{d}t} = -\vec{S}_n \times \frac{\partial H}{\partial \vec{S}_n} \quad . \tag{1.4}$$

Obviously, this can be obtained also by computing $\mathrm{d}S_n^i/\mathrm{d}t = \{S_n^i, H\}$.

Finally, substituting H into the equations of motion we get

$$\frac{\mathrm{d}\vec{S}_n}{\mathrm{d}t} = \vec{S}_n \times \left[\vec{h}_n + \sum_m I_{nm}\vec{S}_m\right] \quad . \tag{1.5}$$

1.1.2 Isotropic Heisenberg chain and XXZ chain

When the interaction among the spin in a 1D chain is isotropic in the sense previously mentioned, $I_{nm} = g_{nm}\mathbb{1}_3$, and it concerns only first neighbors, the model is called **isotropic Heisenberg chain** or **XXX chain**; when the interaction is isotropic only on the x and y directions, i.e.

$$I_{nm} = g_{nm}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix} \quad , \tag{1.6}$$

the model is called **XXZ chain**. When in presence of constant external magnetic field $\vec{h}$ it is possible in both cases to perform a simple change of variables — a time dependent rotation, precisely — that can get rid for us of the contribution of the external field to the Hamiltonian. In the case of the fully isotropic model the operation is always allowed, while for the XXZ chain we must require that the direction of anisotropy is the same of that of the external field, i.e. $\vec{h}$ goes along z.

Removal of constant external field

The Hamiltonian is

$$H = -\frac{1}{2}\sum_{n,m} \vec{S}_n \cdot I_{nm}\vec{S}_m - \sum_n \vec{h}_n \cdot \vec{S}_n \quad , \tag{1.7}$$

with the equations of motion

$$\frac{\mathrm{d}\vec{S}_n}{\mathrm{d}t} = \vec{S}_n \times \vec{h}_n + \vec{S}_n \times \sum_m I_{nm}\vec{S}_m \quad . \tag{1.8}$$

Let us define the time dependent transformation $e^{tA} \in \mathrm{SO}(3)$

$$\vec{S}_n = e^{tA}\vec{\sigma}_n \quad , \tag{1.9}$$

such that, for any vector $\vec{v}$,

$$A\vec{v} = -\vec{h} \times \vec{v} \quad . \tag{1.10}$$

In practice, this means that

$$A = \begin{bmatrix} 0 & h_3 & -h_2 \\ -h_3 & 0 & h_1 \\ h_2 & -h_1 & 0 \end{bmatrix} \quad , \tag{1.11}$$

so the matrix e^{tA} describes a uniform rotation around the axis $\vec{h}$, as e^{tA} is the flow of the equation $\mathrm{d}\vec{S}/\mathrm{d}t = \vec{S} \times \vec{h}$.

We will now insert the transformation into our equations of motion. Before we do that, let us mention that for any rotation matrix R and any pair of vectors $\vec{u}$, $\vec{v}$, the following identity holds:

$$R(\vec{u} \times \vec{v}) = (R\vec{u}) \times (R\vec{v}) \quad . \tag{1.12}$$

That said, whenever the rotation matrix commutes with the interaction matrix, we find:

$$e^{tA}\frac{d\vec{\sigma}_n}{dt} = (e^{tA}\vec{\sigma}_n) \times (e^{tA}\sum_m I_{nm}\vec{\sigma}_m) \quad , \tag{1.13}$$

so that

$$\frac{d\vec{\sigma}_n}{dt} = \vec{\sigma}_n \times \sum_m I_{nm}\vec{\sigma}_m \quad , \tag{1.14}$$

But these are the equations of motion for the usual Hamiltonian

$$H = -\frac{1}{2}\sum_n \vec{\sigma}_n \cdot I_{nm}\vec{\sigma}_m \quad , \tag{1.15}$$

so the external field has been removed.

This is always true for the isotropic case as the interaction is basically only the identity. In the case of the XXZ chain, since the rotation is along $\vec{h}$, the rotation matrix e^{tA} commutes with the interaction matrix (1.6) when $\vec{h}$ itself is along z, and this is the only case in which we can formally eliminate the contribution of the external field. Notice that this operation can be performed under these conditions on a spin lattice of arbitrary dimension, not just one-dimensional.

From now on, we will assume the conditions of a 1D Heisenberg model with constant external field $\vec{h}$ along z — and we will continue the discussion with only the spin-spin interaction term.

Ferromagnetic XXX chain

Let us look at what remains of the Hamiltonian in the general form (1.1) when we make the assumption of XXX chain. Here we have isotropic and nearest neighbor interaction, i.e.

$$I_{nm} = \begin{cases} g\mathbb{1}_3 \ , \quad g > 0 \ , \text{ if } |n-m| = 1 \\ 0 \quad \text{otherwise} \end{cases} \quad , \tag{1.16}$$

and plugging it into the Hamiltonian gives us

$$H = -\frac{g}{2} \sum_{n=1}^{N} \vec{S}_n \cdot \left(\vec{S}_{n+1} + \vec{S}_{n-1} \right) = -g \sum_{n=1}^{N} \vec{S}_n \cdot \vec{S}_{n+1} \tag{1.17}$$

where the equations of motion are

$$\frac{\mathrm{d}\vec{S}_n}{\mathrm{d}t} = g \ \vec{S}_n \times \left(\vec{S}_{n+1} + \vec{S}_{n-1} \right) \quad . \tag{1.18}$$

In this model, the ferromagnetic behavior is determined by the condition $g > 0$. Indeed, from a thermodynamical point of view, the energy is minimized in the situation of all spins aligned, the minimum value of H being $-gN$ on any equilibrium configuration $\vec{S}_n = \vec{S} \ , \ \forall n = 1, ..., N$.

1.2 Continuum limit

Let us now perform the continuum limit for the case of the XXX chain. Doing so will lead us to an important, well known equation: the *Landau-Lifshitz equation*.

We imagine that there exists a function which interpolates our values of the spins in space and time, meaning

$$\begin{aligned} \vec{S} : & \mathbb{R} \times \mathbb{T} \longrightarrow S_2 \\ & (T, x) \longmapsto \vec{S}(T, x) \quad , \end{aligned} \tag{1.19}$$

such that

$$\vec{S}_n(t) = \vec{S}(T, x)\big|_{\substack{T=t/N \\ x=n/N}} \quad . \tag{1.20}$$

Using the new map, the equations of motion become

$$\frac{\partial \vec{S}}{\partial T} = \frac{g}{N} \left(\vec{S} \times \frac{\partial^2 \vec{S}}{\partial x^2} + O(\frac{1}{N^2}) \right) \Bigg|_{\substack{T=t/N \\ x=n/N}} \quad , \tag{1.21}$$

while the Hamiltonian is

$$H = -\frac{1}{2}\frac{g}{N}\sum_n \frac{1}{N}\left(\vec{S}\cdot\frac{\partial^2\vec{S}}{\partial x^2} + O(\frac{1}{N^2})\right)\Bigg|_{\substack{T=t/N \\ x=n/N}} - \frac{g}{N} \quad . \tag{1.22}$$

Given the appearance of the ratio g/N, it is natural assume the existence of

$$\lim_{N\to\infty}\frac{g}{N} := \gamma \quad . \tag{1.23}$$

Upon performing the actual limit $N \to \infty$ we find

$$\mathcal{H} := \lim_{N\to\infty}(H + gN) = -\frac{1}{2}\gamma\int_{\mathbb{T}}\vec{S}\cdot\frac{\partial^2\vec{S}}{\partial x^2}\mathrm{d}x = \frac{\gamma}{2}\int_{\mathbb{T}}\left|\frac{\partial\vec{S}}{\partial x}\right|^2\mathrm{d}x \quad , \tag{1.24}$$

with equations of motion

$$\frac{\partial\vec{S}}{\partial T} = \gamma\vec{S}\times\frac{\partial^2\vec{S}}{\partial x^2} = -\vec{S}\times\frac{\delta\mathcal{H}}{\delta\vec{S}} \quad . \tag{1.25}$$

This PDE, which we rewrite in the compact form

$$\partial_T\vec{S} = \gamma\vec{S}\times\vec{S}_{xx} \quad , \tag{1.26}$$

is precisely the so called *Landau-Lifshitz equation*, which, along with its variations, is the central equation for describing the dynamics of magnetism and micromagnetism [1].

1.2.1 Linear waves

It is interesting to see how the continuum limit for the spin chain, perturbed around an equilibrium solution, is mapped into the Schr'odinger equation for the free particle. One can see that any constant vector $\vec{S}_0$ is a solution to the Landau-Lifshitz equation (1.26); let us perturb such a solution, where for example $\vec{S}_0$ is the unit vector in the z direction (for simplicity we can assume to have unit spin and $\gamma = 1$ without having to change our discussion up to now). The spin function is then $\vec{S} = \hat{e}_z + \vec{\sigma}$, and if we linearize the resulting equation in $\vec{\sigma}$, we find

$$\frac{\partial\vec{\sigma}}{\partial T} = \hat{e}_z\times\partial_x^2\vec{\sigma} \quad \Longleftrightarrow \quad \begin{cases}\partial_T\sigma_x = -\partial_x^2\sigma_y \\ \partial_T\sigma_y = \partial_x^2\sigma_x\end{cases} \quad . \tag{1.27}$$

This is a Hamiltonian system since the proper Hamiltonian (so that the Hamilton equations for σ_x and σ_y are satistisfied) can be found in the following:

$$\mathcal{H} := \frac{1}{2}\int_{\mathbb{T}}[(\partial_x\sigma_x)^2 + (\partial_x\sigma_y)^2]\mathrm{d}x \quad . \tag{1.28}$$

We can now define the complex field

$$\psi := \frac{\sigma_x + i\sigma_y}{\sqrt{2}} \quad , \quad \psi^* := \frac{\sigma_x - i\sigma_y}{\sqrt{2}} \quad , \tag{1.29}$$

which transforms the usual symplectic Poisson tensor $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ into $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$. Its Hamiltonian reads

$$\mathcal{H} = \int_{\mathbb{T}} |\partial_x\psi|^2\mathrm{d}x \tag{1.30}$$

with associated equations of motion

$$\partial_T\psi = -i\frac{\delta\mathcal{H}}{\delta\psi^*} = i\partial_x^2\psi \tag{1.31}$$

and its complex conjugate, namely the Schr$\acute{\text{i}}$odinger equation for the free particle on $\mathbb{T}$. Moreover, if we look for plane-wave solutions of the form

$$\psi(T,x) = Ae^{i(kx-\omega T)} \quad , \tag{1.32}$$

we find the dispersion relation $\omega = k^2$, with $k = 2\pi j$, $j \in \mathbb{Z}$.

Chapter 2

The quantum model

The classical formalism for spin chains is certainly interesting and intuitive, and has been studied at great length for this reason. Nevertheless, it is not quite an accurate description for the phenomena it wishes to study, as the concept of spin and its consequences are an intrinsic feature of matter at a quantum level. The quantum models of spin systems, in fact, revert to their classical correspondents in the limit $S \to \infty$, which is very far from our $S = 1/2$. Because of this, we will go through the process of quantizing the Heisenberg chain in order to later work upon its quantum version. This will require, naturally, to move to the quantum notion of spin in terms of spin operators; after that, we will be able to act somewhat freely as to which variables and operators we wish to use to describe our system.

2.1 First quantization

We begin the process of quantizing the Heisenberg isotropic spin 1/2 chain by studying the model in a first quantization work frame, promoting the physical quantities from functions to operators. We will see that, once we go over the technicalities of working with operators, the model remains formally similar to its classical counterpart.

2.1.1 Heisenberg ferromagnetic chain in first quantization formalism

In the Heisenberg model of the classical spin chain, the prime quantities playing a role are the spin vector components. The spin is described by representations of the SU(2) Lie group and acts upon a $(2s+1)$-dimensional Hilbert space. In our case of spin 1/2, then, the spin component operators

on each site are described by half the Pauli matrices (remember that we are working with the convention $\hbar = 1$), and the commutation relations for such operators are, naturally,

$$[\hat{S}_m^i, \hat{S}_n^j] = i\delta_{mn}\epsilon_{ijk}\hat{S}_n^k \quad . \tag{2.1}$$

We now proceed by substituting the classical components with the new operators in the Heisenberg Hamiltonian, which becomes

$$\begin{aligned} \hat{H} &= -\vec{h} \cdot \sum_n \hat{\vec{S}_n} - \frac{1}{2}\sum_n g\Big(\hat{\vec{S}_n} \cdot \hat{\vec{S_{n+1}}} + \hat{\vec{S}_n} \cdot \hat{\vec{S_{n-1}}}\Big) \\ &= -\vec{h} \cdot \sum_n \hat{\vec{S}_n} - g\sum_n \hat{\vec{S}_n} \cdot \hat{\vec{S_{n+1}}} \quad , \end{aligned} \tag{2.2}$$

as any component commutes with the others when on different sites. We start to notice that, visually speaking, there is not much difference between this quantized Hamiltonian and the classical one. We will see that, in this formalism, the similarity persists further in our analysis of the discrete case.

Equations of motion

Prior to calculating the equations of motion, it is useful to make the following observation.

Observation 2.1.1. *For any vector $\vec{v}$, we have*

$$[\hat{\vec{S}_n}, \vec{v} \cdot \hat{\vec{S}_n}] = i(\vec{v} \times \hat{\vec{S}_n}) \quad . \tag{2.3}$$

Proof.

$$\begin{aligned} [\hat{S}_n^x, \vec{v} \cdot \hat{\vec{S}_n}] &= [\hat{S}_n^x, v^x\hat{S}_n^x + v^y\hat{S}_n^y + v^z\hat{S}_n^z] = \\ &= [\hat{S}_n^x, v^y\hat{S}_n^y + v^z\hat{S}_n^z] = i(v^y\hat{S}_n^z - v^z\hat{S}_n^y) \quad , \end{aligned}$$

and so on for the other components. □

In the eyes of the operator $\hat{\vec{S}_n}$, any other spin $\hat{\vec{S}_m}$ with $n \neq m$ has no more meaning than a mere vector, since any two components on different sites commute. In light of this we see that we can extensively make use of observation (2.1.1) in the computation of the equations of motion, just so:

$$\frac{d\hat{\vec{S}_n}}{dt} = -i[\hat{\vec{S}_n}, \hat{H}]$$

$$= i[\hat{\vec{S}}_n, \vec{h} \cdot \hat{\vec{S}}_n] + ig[\hat{\vec{S}}_n, \hat{\vec{S}}_n \cdot \hat{\vec{S}}_{n+1}] + ig[\hat{\vec{S}}_n, \hat{\vec{S}}_n \cdot \hat{\vec{S}}_{n-1}] =$$
$$= i(i\vec{h} \times \hat{\vec{S}}_n) + ig(i\hat{\vec{S}}_{n+1} \times \hat{\vec{S}}_n) + ig(i\hat{\vec{S}}_{n-1} \times \hat{\vec{S}}_n) =$$
$$= \hat{\vec{S}}_n \times \left[\vec{h} + g(\hat{\vec{S}}_{n-1} + \hat{\vec{S}}_{n+1})\right] \quad . \tag{2.4}$$

These are, as said before, completely analogous to the classical Heisenberg model equations of motion.

2.1.2 Elimination of constant external field

It is easy to see that the removal of the external field $\vec{h}$ in the case of first quantization is conceptually the same procedure as in the classical case, with only a few more details to review. Indeed it has to be noted that the time dependent SO(3) transformation we previously applied in 1.1.2 of the classical case will now only affect the "vectorial" nature of our quantities, leaving their "operator" feature — that is, the real new feature of this approach — untouched.

The only concern we have now is that the new rotated spin, this time an operator, $\hat{\vec{\sigma}}_n$, remains a spin in the quantum sense, i.e. that it keeps the usual commutation relations. This fact is intuitively true, and very easy to check. Naturally, one can see that the commutation relations for the spin can be encoded in the following:

$$\hat{\vec{S}}_n \times \hat{\vec{S}}_n = i\hat{\vec{S}}_n \quad . \tag{2.5}$$

It is then obvious that applying a rotation to the vector $\hat{\vec{S}}_n$ will not change this identity, thanks to the already mentioned properties of SO(3).

Let us now briefly retrace the steps of field removal. First, we define the transformation $e^{tA} \in \mathrm{SO}(3)$

$$\hat{\vec{S}}_n = e^{tA}\hat{\vec{\sigma}}_n \quad , \tag{2.6}$$

such that, for any vector $\vec{v}$,

$$A\vec{v} = -\vec{h} \times \vec{v} \quad . \tag{2.7}$$

Then we use it manifestly in the equations of motion obtained earlier and we find

$$\frac{d\hat{\vec{\sigma}}_n}{dt} = g\hat{\vec{\sigma}}_n \times (\hat{\vec{\sigma}}_{n-1} + \hat{\vec{\sigma}}_{n+1}) \quad . \tag{2.8}$$

We successfully managed once again to eliminate the external field, as these are the equations of motion for the usual Hamiltonian

$$\hat{H} = -g \sum_n \hat{\vec{\sigma}}_n \cdot \hat{\vec{\sigma}}_{n+1} \quad . \tag{2.9}$$

Just like in the classical case, we will henceforth continue our dissertation considering no external field $\vec{h}$ without loss of generality, since we can always revert back to the original system with a rotation of the spins.

Going forward

To go forth with the analogy between classical and first quantized system one would begin to perform the continuum limit in order to obtain a first quantized version of the Landau-Lifshitz equation. We leave it to the reader to check that the process of doing so is formally the same, leading simply to the equation

$$\partial_T \hat{\vec{S}} = \hat{\vec{S}} \times \partial_x^2 \hat{\vec{S}} \quad . \tag{2.10}$$

2.2 Ladder operators

As we saw, there is really nothing interesting about a first quantized version of the Heisenberg chain — the discussion follows step by step its classical version, with no additional information to be extracted from the quantum model. It is more interesting, both for technical and conceptual reasons, to resort to ladder operators. From a technical point of view, this is an in-between step which is necessary in order to later shift to a second quantization approach (we will map the Heisenberg model into a system of particles). From a conceptual one, there is a great deal more to say about the system when using such operators; the famous *Bethe Ansatz* [11], the first historical solution to the model and the most vastly researched, was made studying the system precisely in the work frame of ladder operators.

The new set of operators in play is then made up of the following two:

$$\hat{S}_n^+ := \hat{S}_n^x + i\hat{S}_n^y \quad , \tag{2.11}$$

$$\hat{S}_n^- := \hat{S}_n^x - i\hat{S}_n^y \quad , \tag{2.12}$$

along with $\hat{S}^z$, which stays the same. Every operator can now act on states like $|\uparrow_1, \uparrow_2, \downarrow_3, ..., \uparrow_N\rangle$ which symbolize the spin up-down configuration of the

system. The action of the ladder operators on such states is

$$\begin{aligned}
&\hat{S}_n^+ \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \uparrow_n, ..., \uparrow_N\right\rangle = 0 \quad ; \\
&\hat{S}_n^+ \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \downarrow_n, ..., \uparrow_N\right\rangle = \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \uparrow_n, ..., \uparrow_N\right\rangle \quad ; \qquad (2.13) \\
&\hat{S}_n^- \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \uparrow_n, ..., \uparrow_N\right\rangle = \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \downarrow_n, ..., \uparrow_N\right\rangle \quad ; \\
&\hat{S}_n^- \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \downarrow_n, ..., \uparrow_N\right\rangle = 0 \quad ; \qquad (2.14)
\end{aligned}$$

while $\hat{S}^z$ acts like

$$\begin{aligned}
&\hat{S}_n^z \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \uparrow_n, ..., \uparrow_N\right\rangle = \frac{1}{2} \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \uparrow_n, ..., \uparrow_N\right\rangle \quad ; \\
&\hat{S}_n^z \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \downarrow_n, ..., \uparrow_N\right\rangle = -\frac{1}{2} \left|\uparrow_1, \uparrow_2, \downarrow_3, ..., \downarrow_n, ..., \uparrow_N\right\rangle \quad ; \qquad (2.15)
\end{aligned}$$

2.2.1 Heisenberg in ladder operators

From the previous Section, we know that the Heisenberg Hamiltonian reads

$$\hat{H} = -g \sum_n \hat{S}_n \cdot \hat{S}_{n+1} \quad , \qquad (2.16)$$

where the commutation relations for the $\hat{S}_n$ operators are

$$[\hat{S}_m^i, \hat{S}_n^j] = i\delta_{mn}\epsilon_{ijk}\hat{S}_n^k \quad . \qquad (2.17)$$

From these, we find that the commutation relations for the new operators are

$$[\hat{S}^+, \hat{S}^-] = 2\hat{S}^z \quad , \quad [\hat{S}^z, \hat{S}^+] = \hat{S}^+ \quad , \quad [\hat{S}^z, \hat{S}^-] = -\hat{S}^- \quad . \qquad (2.18)$$

Just from the definition of $\hat{S}_n^+$ and $\hat{S}_n^-$, we know that they deal only with the x and y directions. It is then straightforward to check that the x-y interaction term in the Hamiltonian can be written in terms of ladder operators simply via the following identity:

$$\hat{S}_n^+ \hat{S}_{n+1}^- + \hat{S}_n^- \hat{S}_{n+1}^+ = 2\Big[\hat{S}_n^x \hat{S}_{n+1}^x + \hat{S}_n^y \hat{S}_{n+1}^y\Big] \quad . \qquad (2.19)$$

Thus, since the interaction along z remains unaltered, we find the new Hamiltonian to be

$$\hat{H} = -g \sum_n \left(\frac{1}{2}\Big[\hat{S}_n^+ \hat{S}_{n+1}^- + \hat{S}_n^- \hat{S}_{n+1}^+\Big] + \hat{S}_n^z \hat{S}_{n+1}^z\right) \quad . \qquad (2.20)$$

2.2.2 Towards another change of coordinates

It is interesting to notice that there is an abundance of operators in use, i.e. there exists an identity which relates $\hat{S}_n^+$, $\hat{S}_n^-$ and $\hat{S}_n^z$. This follows from the fact that the square of the spin on each site has just one eigenvalue, $s(s+1) = 3/4$, so

$$\frac{3}{4} = \frac{1}{2}\left[\hat{S}_n^+\hat{S}_n^- + \hat{S}_n^-\hat{S}_n^+\right] + (\hat{S}_n^z)^2 \quad . \tag{2.21}$$

Therefore, we are allowed to look for a change of variables that, for each site n, maps our three spin-related operators into two new operators.

It is also worth noticing that the raising and lowering operators here introduced possess a quite peculiar property, and that is what will finally enable us to make the crucial change of variables. Such property is simply

$$(S_n^+)^2 = (S_n^-)^2 = 0 \quad , \tag{2.22}$$

$$S_n^- S_n^+ + S_n^+ S_n^- = 1 \quad . \tag{2.23}$$

Clearly, this already looks like some fermionic creation and annihilation operator property. As we will see in Section 2.3, it is thanks to this that we will be able to transfer precisely into a fermionic Fock space workframe, via the so called *Jordan-Wigner transformation*.

2.3 Jordan-Wigner transformation

As we already mentioned in section 2.2, equations (2.22) and (2.23) are similar to the relations pertaining the fermion creation and annihilation operators *on one site*. Meaning,

$$(f_n^\dagger)^2 = (f_n)^2 = 0 \quad , \tag{2.24}$$

$$\{f_n, f_n^\dagger\} = f_n f_n^\dagger + f_n^\dagger f_n = 1 \quad . \tag{2.25}$$

Unfortunately, this similarity does not hold when we take into account spin operators on different sites; such spin operators clearly commute when on different sites, while generic fermion will keep the anti-commutation relations

$$\{f_n^\dagger, f_m^\dagger\} = \{f_n, f_m\} = 0 \quad , \quad \{f_n, f_m^\dagger\} = \delta_{nm} \quad . \tag{2.26}$$

This means that we need to be careful if we want to properly map a spin chain into a fermion system; we cannot simply replace the spin raising and lowering operators with the correspondent fermionic creation and annihilation ones.

The correct (although longer) procedure has been brought to us in 1928 by Pascual Jordan and Eugene Wigner, opening a portal between spin-1/2 spin chains and fermion chains - spinless fermions, to be exact. Reference an notes on the Jordan-Wigner transformation are [15], [23], [14], [24]. The change of variable is the following:

$$\begin{aligned} S_n^- &= Q_n f_n = \exp\left\{ i\pi \sum_{l=1}^{n-1} f_l^\dagger f_l \right\} f_n \quad , \\ S_n^+ &= f_n^\dagger Q_n^\dagger = f_n^\dagger \exp\left\{ -i\pi \sum_{l=1}^{n-1} f_l^\dagger f_l \right\} \quad , \\ S_n^z &= f_n^\dagger f_n - 1/2 \quad . \end{aligned} \tag{2.27}$$

Here the factor Q_n, which we will see is self-adjoint, is the key which lets our spin operators commute on different sites while the operators f_n keep on being fermions. For practical reasons and from now on, let us use a different form of this factor, as presented in the following lemma:

Lemma 2.3.1.

$$Q_n = \exp\left\{ i\pi \sum_{l=1}^{n-1} f_l^\dagger f_l \right\} = \prod_{l=1}^{n-1} (1 - 2 f_l^\dagger f_l) \quad . \tag{2.28}$$

Proof. Let us notice that, for any integer l, we have that

$$(f_n^\dagger f_n)^l = f_n^\dagger f_n \quad , \quad (f_n^\dagger f_n)^0 = 1 \quad .$$

Then, clearly,

$$\begin{aligned} e^{i\pi f_l^\dagger f_l} = \sum_{n=0}^{\infty} \frac{(i\pi)^n}{n!} (f_l^\dagger f_l)^n = 1 + \left(\sum_{n=1}^{\infty} \frac{(i\pi)^n}{n!} \right) f_l^\dagger f_l = \\ = 1 + (e^{i\pi} - 1) f_l^\dagger f_l = 1 - 2 f_l^\dagger f_l \quad . \end{aligned}$$

The same exact calculation holds for $Q_n^\dagger$, showing that $Q_n^\dagger = Q_n$. □

Let us quickly interpret this map in terms of Hilbert spaces. From the change of variables (2.27), we can see that the fermion operators act on the standard spin basis in such a way:

$$\begin{aligned} f_n \left|\uparrow\right\rangle &= \left|\downarrow\right\rangle \ , \quad f_n \left|\downarrow\right\rangle = 0 \ , \\ f_n^\dagger \left|\downarrow\right\rangle &= \left|\uparrow\right\rangle \ , \quad f_n^\dagger \left|\uparrow\right\rangle = 0 \ . \end{aligned} \tag{2.29}$$

Therefore, by defining the particle number operator $n_n = f_n^\dagger f_n$, it is intuitive that the downward spin state corresponds to an empty site in the fermionic chain, and that the upward spin state means an occupied site in the fermionic chain:

$$n_n \left|\uparrow\right\rangle = \left|\uparrow\right\rangle \ , \quad n_n \left|\downarrow\right\rangle = 0 \ . \tag{2.30}$$

It is clear now why the fermions here are regarded as spinless; the only two possible states for each fermion are "present" and "not present", i.e. a fermion number n_n that can be only 0 or 1. This means that our particles are, indeed, fermions, as they obey the same commutation relations, yet they possess no internal degree of freedom, not even spin.

The anti-commutation property of fermions and the nature of the factor Q_n give birth to many other not-so-intuitive relations, which will reveal crucial for the computation of our fermion Hamiltonian. We will present them here, much like a small handbook of properties:

$$\begin{aligned}
&f_n(1-2f_n^\dagger f_n) = -f_n \quad , \quad f_n^\dagger(1-2f_n^\dagger f_n) = f_n^\dagger \ , \\
&(1-2f_n^\dagger f_n)f_n = f_n \quad , \quad (1-2f_n^\dagger f_n)f_n^\dagger = -f_n^\dagger \ , \\
&(1-2f_n^\dagger f_n)^2 = 1 \ , \\
&f_n(1-2f_l^\dagger f_l) = (1-2f_l^\dagger f_l)f_n \ , \ l \neq n \ , \\
&f_n^\dagger(1-2f_l^\dagger f_l) = (1-2f_l^\dagger f_l)f_n^\dagger \ , \ l \neq n \ , \\
&f_n Q_l = Q_l f_n \ , \ l \leq n \ , \\
&f_n^\dagger Q_l = Q_l f_n^\dagger \ , \ l \leq n \ , \\
&Q_n Q_{n+1} = 1 - 2f_n^\dagger f_n \ , \\
&Q_n^2 = 1 \ .
\end{aligned} \tag{2.31}$$

Reverse Jordan-Wigner

For the sake of completeness we report here the reverse identities for the Jordan-Wigner transformation, i.e. the expression of the fermion operators in terms of spin, obtained using equations (2.31). They are:

$$\begin{aligned}
f_n &= \prod_{l=1}^{n-1}(-2S_l^z)S_n^- \quad , \\
f_n^\dagger &= S_n^+ \prod_{l=1}^{n-1}(-2S_l^z) \quad , \\
n_n &= S_n^z + 1/2 \quad .
\end{aligned} \tag{2.32}$$

2.4 The Heisenberg fermion system

Now that we have seen the technicalities of the Jordan-Wigner transformation, we may proceed to turn our Heisenberg spin chain into a system of fermions located on a discrete number of sites in one dimension. This has been done extensively starting from the XY model instead of the full Heisenberg model, with the result, as we will see for ourselves, of a mapping onto a system of free fermions. Applying the transformation on the full model, complete with interaction among the z-components, will naturally provide us with additional contributions in the fermion Hamiltonian, i.e. with interaction.

2.4.1 Transformed Heisenberg Hamiltonian

We recall from the previous section the following change of coordinates:

$$\hat{S}_n^- = Q_n f_n \tag{2.33}$$

$$\hat{S}_n^+ = f_n^\dagger Q_n \tag{2.34}$$

$$\hat{S}_n^z = f_n^\dagger f_n - 1/2 \quad . \tag{2.35}$$

By the properties of these new operators and coefficients widely explored in the same section, the key components of our Hamiltonian behave in such a way:

$$\begin{aligned} \hat{S}_n^+ \hat{S}_{n+1}^- &= f_n^\dagger Q_n Q_{n+1} f_{n+1} = f_n^\dagger (1 - 2 f_n^\dagger f_n) f_{n+1} = \\ &= f_n^\dagger f_{n+1} \quad ; \end{aligned} \tag{2.36}$$

$$\begin{aligned} \hat{S}_n^- \hat{S}_{n+1}^+ &= f_n Q_n Q_{n+1} f_{n+1}^\dagger = f_n (1 - 2 f_n^\dagger f_n) f_{n+1}^\dagger = \\ &= -f_n f_{n+1}^\dagger = f_{n+1}^\dagger f_n \quad . \end{aligned} \tag{2.37}$$

Now we clearly have all the elements to bring the Heisenberg spin chain to a fermion system. Before rushing into it, though, let us slowly study the Hamiltonian piece by piece; meaning, let us first observe how the x-y interaction alone (i.e. the XY model) is mapped into a Hamiltonian which is quadratic in the fermion operators and that represents a free system of fermions. Let us also study the contribution of the $\hat{S}^z$ interaction alone, and see how it provides us with a quartic term (an interaction term) as well as another quadratic term that can be put together with the XY terms to make up the integrable part of our system. But we're getting ahead of ourselves.

The XY model

The standard XY model Hamiltonian has the same form as the Heisenberg one, but with the interaction - once again homogeneous - only present in the x and y directions. Namely,

$$\hat{H}_{XY} = -g \sum_n \left[\hat{S}^x_n \hat{S}^x_{n+1} + \hat{S}^y_n \hat{S}^y_{n+1} \right] = \\ = -\frac{1}{2} g \sum_n \left[\hat{S}^+_n \hat{S}^-_{n+1} + \hat{S}^-_n \hat{S}^+_{n+1} \right] \quad . \tag{2.38}$$

From equations (2.37) it is immediate to see that the Hamiltonian in terms of the fermionic operators reads

$$\hat{H}_{XY} = -\frac{1}{2} g \sum_n \left[f^\dagger_n f_{n+1} + f^\dagger_n f_{n-1} \right] \quad . \tag{2.39}$$

This is not a free fermion Hamiltonian at first sight. However, any Hamiltonian of the form $H = \sum_{jk} \alpha_{jk} f^\dagger_j f_k$ can be converted into a free fermion Hamiltonian by means of a simple transformation. Our case clearly fits in this family, as it is just the situation in which $\alpha_{jk} = -\frac{1}{2} g(\delta_{k,j+1} + \delta_{k,j-1})$. The procedure is fully explained in [25] for the case of a more general family of Hamiltonians which are quadratic in the Fermi operators. Such generality is not of particular interest for us, so we only report here a simpler version of the transformation (in particular case, this will take the form of a discrete Fourier transform).

Keeping in mind that for the Hamiltonian to be hermitian we would require $\alpha^*_{jk} = \alpha_{kj}$ (i.e. the matrix α is hermitian itself), we introduce the operators

$$\varphi_j := \sum_k u_{jk} f_k \quad , \quad u_{jk} \in \mathbb{C} \quad , \tag{2.40}$$

where we want the u_{jk} to be such that the commutation relations are preserved (i.e. that the φ_j are still fermions). Namely

$$\{\varphi_j, \varphi_k\} = \sum_{lm} u_{jl} u^*_{km} \{f_l, f_m\} = \sum_{lm} u_{jl} \delta_{lm} u^\dagger_{mk} \mathbb{1} = \\ = (uu^\dagger)_{jk} \mathbb{1} \quad , \tag{2.41}$$

so that we require $uu^\dagger$ to be the identity, i.e. that u be a unitary transformation. Given this new information, we can then revert the relation between the two fermion operators, obtaining

$$f_j := \sum_k u^\dagger_{jk} \varphi_k \quad , \tag{2.42}$$

and insert it into H. In this way, the Hamiltonian has the form

$$H = \sum_{jk} \alpha_{jk} f_j^\dagger f_k = \sum_{jk} \sum_{lm} \alpha_{jk} \varphi_l^\dagger u_{lj} u_{km}^\dagger \varphi_m = \tag{2.43}$$

$$= \sum_{lm} \varphi_l^\dagger \Big(\sum_{jk} u_{lj} \alpha_{jk} u_{km}^\dagger \Big) \varphi_m = \sum_{lm} \varphi_l^\dagger \Big(u \alpha u^\dagger \Big)_{lm} \varphi_m \quad . \tag{2.44}$$

Since α is hermitian, we can still choose the transformation u such that it diagonalizes it. If we let λ_j be the eigenvalues of α, we then have

$$H = \sum_j \lambda_j \varphi_j^\dagger \varphi_j \quad , \tag{2.45}$$

which is manifestly the Hamiltonian for a free system of fermions.

The Heisenberg Model

Adding the $\hat{S}^z$ interaction to the XY model and substituting S^z with the correspondent according to the Jordan-Wigner, we get the fermionic formulation of the Heisenberg model, i.e.

$$\begin{aligned}
\hat{H} &= -g \sum_n \left(\frac{1}{2} \Big[\hat{S}_n^+ \hat{S}_{n+1}^- + \hat{S}_n^- \hat{S}_{n+1}^+ \Big] + \hat{S}_n^z \hat{S}_{n+1}^z \right) \\
&= -g \sum_n \left[\Big(f_n^\dagger f_n - \frac{1}{2} \Big) \Big(f_{n+1}^\dagger f_{n+1} - \frac{1}{2} \Big) + \frac{1}{2} \Big(f_n^\dagger f_{n+1} + f_n^\dagger f_{n-1} \Big) \right] \\
&= -g \sum_n \left[\frac{1}{4} + \frac{1}{2} \Big(f_n^\dagger f_{n+1} - 2 f_n^\dagger f_n + f_n^\dagger f_{n-1} \Big) + f_n^\dagger f_n f_{n+1}^\dagger f_{n+1} \right] \\
&= \hat{H}_0 + \hat{H}_z \quad ,
\end{aligned} \tag{2.46}$$

where we defined $\hat{H}_0$ and $\hat{H}_z$ so that the Hamiltonian is split into a quadratic term (leading term) and a quartic term of interaction that will be later treated as a perturbation. Namely,

$$\hat{H}_0 = -g \sum_n \left[\frac{1}{4} + \frac{1}{2} \Big(f_n^\dagger f_{n+1} - 2 f_n^\dagger f_n + f_n^\dagger f_{n-1} \Big) \right] \quad , \tag{2.47}$$

$$\hat{H}_z = -g \sum_n f_n^\dagger f_n f_{n+1}^\dagger f_{n+1} \quad . \tag{2.48}$$

The unperturbed term, $\hat{H}_0$, is then basically the XY model Hamiltonian plus a quadratic term (which is a sum of number operators). This additional term, reverted to S variables, reads

$$g \sum_n f_n^\dagger f_n = g \sum_n \left[S_n^z - \frac{1}{2} \right] \quad , \tag{2.49}$$

meaning that what we treat as the unperturbed system is an XY model with a constant external magnetic field depending on g.

2.4.2 Equations of motion for fermions

Let us shorten the computation of the equations of motion for the fermionic operators by proving the following relations:

Observation 2.4.1.

$$[f_a, f_b^\dagger f_c] = \delta_{ab} f_c \quad , \tag{2.50}$$
$$[f_a^\dagger, f_b^\dagger f_c] = -\delta_{ac} f_b^\dagger \quad . \tag{2.51}$$

Proof.

$$[f_a, f_b^\dagger f_c] = [f_a, f_b^\dagger] f_c + f_b^\dagger [f_a, f_c] = (\delta_{ab} - 2 f_b^\dagger f_a) f_c + f_b^\dagger (-2 f_c f_a) = \\ = \delta_{ab} f_c - 2 f_b^\dagger \{f_a, f_c\} = \delta_{ab} f_c \quad . \tag{2.52}$$

The proof for (2.51) is completely analogous. □

Knowing these relations and the ones presented in Section 2.3, the computation is quite straightforward. We report it here for the sake of completeness, highlighting the origin (quadratic or quartic part of the Hamiltonian) of each contribution:

$$\frac{\mathrm{d} f_a}{\mathrm{d} t} = -i[f_a, \hat{H}] = -i\Big([f_a, \hat{H}_0] + [f_a, \hat{H}_z]\Big) = A + B \quad ,$$

where

$$A = -i[f_a, \hat{H}_0] = i\frac{g}{2} \sum_n [f_a, f_n^\dagger f_{n+1} - 2 f_n^\dagger f_n + f_n^\dagger f_{n-1}] = \\ = i\frac{g}{2} \sum_n \Big(\delta_{an} f_{n+1} - 2\delta_{an} f_n + \delta_{an} f_{n-1}\Big) = i\frac{g}{2}\Big(f_{a+1} - 2 f_a + f_{a-1}\Big) = \\ = i\frac{g}{2} \Delta_d(f_a) \quad ,$$

$$B = -i[f_a, \hat{H}_z] = ig \sum_n [f_a, f_n^\dagger f_n f_{n+1}^\dagger f_{n+1}] = \\ = ig \sum_n \Big([f_a, f_n^\dagger f_n] f_{n+1}^\dagger f_{n+1} + f_n^\dagger f_n [f_a, f_{n+1}^\dagger f_{n+1}]\Big) = \\ = ig \sum_n \Big(\delta_{an} f_n f_{n+1}^\dagger f_{n+1} + \delta_{a,n+1} f_n^\dagger f_n f_{n+1}\Big) = ig\Big(f_a f_{a+1}^\dagger f_{a+1} + f_{a-1}^\dagger f_{a-1} f_a\Big) =$$

$$= ig\Big(f^\dagger_{a+1}f_{a+1} + f^\dagger_{a-1}f_{a-1}\Big)f_a = ig\Big(f^\dagger_{a+1}f_{a+1} + f^\dagger_{a-1}f_{a-1} - 2f^\dagger_a f_a\Big)f_a =$$
$$= ig\Delta_d(f^\dagger_a f_a)f_a \quad .$$

This way, we finally get:

$$\frac{\mathrm{d}f_a}{\mathrm{d}t} = i\frac{g}{2}\Delta_d(f_a) + ig\Delta_d(f^\dagger_a f_a)f_a \quad . \tag{2.53}$$

The symbol $\Delta_d(\phi_n)\big(:= \phi_{n+1} + \phi_{n-1} - 2\phi_n\big)$ is here used as a sort of discrete Laplacian, as its value suggests.

The equations of motion for the creation operator $f^\dagger$ are computed analogously, step by step. Of course, we find exactly the hermitian conjugate of equation (2.53), namely

$$\frac{\mathrm{d}f^\dagger_a}{\mathrm{d}t} = -i\frac{g}{2}\Delta_d(f^\dagger_a) - igf^\dagger_a\Delta_d(f^\dagger_a f_a) \quad . \tag{2.54}$$

Chapter 3

Perturbative approach

The main aim of this dissertation is discussed in this chapter. As announced, we are interested in verifying the possible advantages of Hamiltonian perturbation theory, in its quantum parallel explained in appendix A, by applying it on the quantum Heisenberg chain. For this purpose, we will work with the fermionized version found in the previous chapter, described by the Hamiltonian

$$\hat{H} = -g \sum_n \left[\frac{1}{4} + \frac{1}{2} \left(f_n^\dagger f_{n+1} - 2 f_n^\dagger f_n + f_n^\dagger f_{n-1} \right) \right] - g \sum_n f_n^\dagger f_n f_{n+1}^\dagger f_{n+1} \quad . \tag{3.1}$$

The quadratic part is diagonalizable, i.e. can be easily solved. This provides us with an simple integrable system to start with. The quartic term will then be our perturbation; working with the isotropic chain, though, restricts the goodness of the approximated approach to the case of few overturned spins, or small excitations. The applicability of the method is further discussed in section 3.4.

Section 3.1 will be dedicated to performing a change of variables in order to diagonalize the quadratic term, while section 3.2 will contain the actual application of the Averaging Principle theorem [21] on the perturbative term to find the first order truncated normal form of the Hamiltonian. We will then comment on the findings by also examining the correction to the energy eigenvalues that follows from adding the perturbative term.

3.1 Diagonalization of quadratic (unperturbed) term

We know that the XY model is integrable. This is because, as we know from 2.4.1, any quadratic femionic Hamiltonian is diagonalizable. So, of course, our quadratic term is integrable as well (this is nothing new — we know that it basically corresponds to some XY model). We will therefore diagonalize the system made of the quadratic term in order to solve it, and to later treat it as the unperturbed Hamiltonian, whereas the perturbation itself will be coming from $\hat{H}_z$ once it is put in the right variables. From now on, we choose to ignore any constant term that will appear in the Hamiltonian.

3.1.1 Fourier transform: quadratic term

The procedure described in subsection 2.4.1 to diagonalize quadratic fermionic Hamiltonians offers a transformation which is left in general terms and is not explicit. Fortunately, our case is easily diagonalized via Fourier. Defining the fourier coefficients f_k (which are operators) such that

$$f_n = \frac{1}{\sqrt{N}} \sum_k f_k e^{ikn} \quad , \quad k = \frac{2\pi j}{N} \ , \ j = 1, ..., N \tag{3.2}$$

and remembering that

$$\delta_{kk'} = \frac{1}{N} \sum_n e^{i(k-k')n} \quad , \tag{3.3}$$

we find the free fermion Hamiltonian:

$$\begin{aligned}
\hat{H}_0 &= -g \sum_n \Big(\frac{1}{2} f_n^\dagger f_{n+1} + \frac{1}{2} f_n^\dagger f_{n-1} - f_n^\dagger f_n \Big) = \\
&= -\frac{g}{N} \sum_n \sum_{kk'} \Big[\frac{1}{2} f_k^\dagger f_{k'} e^{-i(k-k')n} \big(e^{ik'} + e^{-ik'} \big) - f_k^\dagger f_{k'} e^{-i(k-k')n} \Big] = \\
&= -g \sum_{kk'} f_k^\dagger f_{k'} (\cos(k') - 1) \delta_{kk'} = g \sum_k (1 - \cos k) f_k^\dagger f_k = \\
&= \sum_k \omega(k) f_k^\dagger f_k \quad ,
\end{aligned} \tag{3.4}$$

where we defined the frequency

$$\omega(k) := g(1 - \cos k) \quad . \tag{3.5}$$

With this precious information we can now solve the dynamics of the XY model. Notice that the commutation relations for the Fourier operators are still exactly the same as for the fermions. To show it very quickly, we recall that expanding in Fourier series can be easily reverted:

$$f_k = \frac{1}{\sqrt{N}} \sum_n f_n e^{-ikn} \quad . \tag{3.6}$$

Therefore,

$$\begin{aligned} \{f_k^\dagger, f_{k'}\} = \frac{1}{N} \sum_{n,m} \{f_n^\dagger, f_m\} e^{ikn} e^{-ik'm} = \frac{1}{N} \sum_{n,m} \delta_{mn} e^{ikn} e^{-ik'm} = \\ = \frac{1}{N} \sum_n e^{i(k-k')n} = \delta_{kk'} \quad , \end{aligned} \tag{3.7}$$

and proceeding similarly for the other anticommutators we get

$$\begin{aligned} \{f_k^\dagger, f_{k'}\} = \{f_k, f_{k'}^\dagger\} = \delta_{kk'} \quad , \\ \{f_k^\dagger, f_{k'}^\dagger\} = \{f_k, f_{k'}\} = 0 \quad . \end{aligned} \tag{3.8}$$

We can then take advantage of the many properties we explored before for fermion operators, which are especially useful when calculating the equations of motion. Explicitly, these are

$$\left. \frac{\mathrm{d} f_q}{\mathrm{d} t} \right|_{H_0} = -i[f_q, \hat{H}_0] = -i \sum_k \omega(k) \delta_{qk} f_k = -i\omega(q) f_q \quad . \tag{3.9}$$

The time evolution of the f_k operators is then simply

$$f_k(t) = f_k e^{-i\omega(k)t} \quad . \tag{3.10}$$

3.1.2 Fourier transform: quartic term

Now that we know the flow (3.10) of the quadratic part of the Hamiltonian alone, we can use this integrable Hamiltonian as the unperturbed part of the system, while treating the quartic part as a perturbation. Clearly, though, we must have the whole system in terms of the same variables, thus we translate into Fourier the quartic term.

$$\hat{H}_z = -g \sum_n f_n^\dagger f_n f_{n+1}^\dagger f_{n+1} =$$

$$
\begin{aligned}
&= - g \sum_n \frac{1}{N^2} \sum k, k', q, q' f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')} e^{-i(k+q-k'-q')n} = \\
&= - g \frac{1}{N} \sum k, k', q, q' f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')} \times \\
&\qquad \times \left(\delta_{k+q,k'+q'} + \delta_{k+q,k'+q'+2\pi} + \delta_{k+q,k'+q'-2\pi} \right) \quad . \qquad (3.11)
\end{aligned}
$$

The sum of deltas here comes from the following:

Lemma 3.1.1. *For $k = \frac{2\pi j}{N}$, $j \in \mathbb{Z}$, the following identity holds:*

$$
\frac{1}{N} \sum_{n=1}^{N} e^{ikn} = \sum_{m \in \mathbb{Z}} \delta_{k,2\pi m} \quad . \qquad (3.12)
$$

Proof. Knowing that $k = 2\pi j/N$, and using the formula for the finite complex geometric series, we find

$$
\frac{1}{N} \sum_{n=1}^{N} e^{ikn} = \frac{1}{N} e^{ik} \sum_{n=0}^{N-1} (e^{ik})^n = \frac{1}{N} e^{ik} \frac{e^{ikN} - 1}{e^{ik} - 1} = \frac{1}{N} e^{i\frac{2\pi j}{N}} \frac{e^{i2\pi j} - 1}{e^{i\frac{2\pi j}{N}} - 1} \quad ,
$$

which is zero whenever $k \neq 2\pi m$, $m \in \mathbb{Z}$. Otherwise, we have

$$
\frac{1}{N} e^{ik} \frac{e^{ikN/2}}{e^{ik/2}} \frac{\sin \frac{kN}{2}}{\sin \frac{k}{2}} = \frac{1}{N} e^{i2\pi m} \frac{(e^{i\pi m})^N}{e^{i\pi m}} \frac{\sin(N\pi m)}{\sin(\pi m)} \quad ,
$$

and using the expansion for the sine of multiple angle

$$
\sin(Nx) = \sum_{\substack{r=0, \\ 2r+1 \leq N}} (-1)^r \begin{pmatrix} N \\ 2r+1 \end{pmatrix} \cos^{N-2r-1}(x) \sin^{2r+1}(x) \quad , \qquad (3.13)
$$

we get

$$
\begin{aligned}
\frac{1}{N} e^{i2\pi m} \frac{(e^{i\pi m})^N}{e^{i\pi m}} \frac{\sin(N\pi m)}{\sin(\pi m)} =& \frac{1}{N} (e^{i\pi m})^{N-1} \sum_{\substack{r=0, \\ 2r+1 \leq N}} (-1)^r \begin{pmatrix} N \\ 2r+1 \end{pmatrix} \times \\
& \times \cos^{N-2r-1}(\pi m) \sin^{2r}(\pi m) = \\
\text{(only } r = 0 \text{ remains)} =& \frac{1}{N} (e^{i\pi m})^{N-1} N (e^{i\pi m})^{N-1} = (-1)^{2m(N-1)} = \\
=& 1 \quad .
\end{aligned}
$$

□

Notice that we already made use of this in (3.3), with the restriction that each k goes from 1 to N. This allows for $k-k'=2\pi m$ the only value $m=0$. It works the same when we have $k-k'+q-q'=2\pi m$; the maximum and minimum values are going to be

$$\max_{k,k',q,q'}(k-k'+q-q') = 4\pi - \frac{4\pi}{N} < 4\pi \tag{3.14}$$

$$\min_{k,k',q,q'}(k-k'+q-q') = -4\pi + \frac{4\pi}{N} > -4\pi \quad , \tag{3.15}$$

so that the only values of m allowed are $m=-1,0,1$.

3.2 Applying perturbation theory

Let us begin the calculation of the first order normal form for our Hamiltonian, following the program described in appendix A. The role of the integrable unperturbed system is played by

$$\hat{h} = \sum_k \omega(k) f_k^\dagger f_k \quad ,$$

whereas the perturbative term, in which we state the time dependence of the operators, is the following:

$$\hat{H}_z = -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger(t) f_{k'}(t) f_q^\dagger(t) f_{q'}(t) e^{-i(q-q')}\Delta_{k+q,k'+q'} \quad , \tag{3.16}$$

where we put

$$\Delta_{k+q,k'+q'} := \delta_{k+q,k'+q'} + \delta_{k+q,k'+q'+2\pi} + \delta_{k+q,k'+q'-2\pi} \quad . \tag{3.17}$$

We proceed by evaluating its time average along the flow (3.10) of $\hat{h}$, $\langle \hat{H}_z \rangle_h$. By doing so, we automatically find the first order perturbative term of the Hamiltonian in normal form, which we call $\hat{S}$.

$$\begin{aligned}
\hat{S} &= \frac{1}{T}\int_0^T \hat{H}_z(t)\Big|_{\text{unperturbed}} \mathrm{d}t \\
&= -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\Delta_{k+q,k'+q'}\frac{1}{T}\int_0^T e^{i[\omega(k)-\omega(k')+\omega(q)-\omega(q')]t}\mathrm{d}t \\
&= -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\Delta_{k+q,k'+q'}\delta_{\omega(k)+\omega(q),\omega(k')+\omega(q')} \quad .
\end{aligned} \tag{3.18}$$

Even though the latter equation could technically constitute the result we are looking for, there is really nothing useful in having it under such appearance; the only information we have added is encoded into the delta of frequencies, but we cannot yet make any use of it. As a matter of fact, to better understand the properties of $\hat{S}$, we will have to solve the system of the two deltas $\Delta_{k+q,k'+q'}$ and $\delta_{\omega(k)+\omega(q),\omega(k')+\omega(q')}$, in order to figure out which combinations of momenta will survive in the sum. The system is pretty straightforward to solve, with only a few remarks to be made, although it is quite a long calculation. The next subsection will be entirely devoted to its illustration.

3.2.1 Solving the delta system

Since $\Delta_{k+q,k'+q'}$ is actually made up of a sum of three deltas, we are in fact dealing with three different systems. Let us solve them one by one, keeping in mind that the systems coming from $\delta_{k+q,k'+q'+2\pi}$ and $\delta_{k+q,k'+q'-2\pi}$ are related by symmetry under the exchange $k \leftrightarrow k'$, $q \leftrightarrow q'$.

First sub-system

Let us start with the symmetric term of the delta:

$$\begin{cases} \omega(k)+\omega(q) = \omega(k')+\omega(q') \\ k+q = k'+q' \end{cases} \implies \begin{cases} \cos(k)+\cos(q) = \cos(k')+\cos(q') \\ k+q = k'+q' \end{cases}$$

$$\begin{cases} \cos(\frac{k+q}{2})\cos(\frac{k-q}{2}) = \cos(\frac{k'+q'}{2})\cos(\frac{k'-q'}{2}) \\ k+q = k'+q' \end{cases} \implies$$

$$\begin{cases} \cos(\frac{k+q}{2})\cos(\frac{k-q}{2}) = \cos(\frac{k+q}{2})\cos(\frac{k'-q'}{2}) \\ k+q = k'+q' \end{cases} \implies$$

$$\begin{cases} \cos(\frac{k+q}{2}) = 0 \\ k+q = k'+q' \end{cases} \vee \begin{cases} \cos(\frac{k-q}{2}) = \cos(\frac{2k'-k-q}{2}) \\ k+q = k'+q' \end{cases}$$

Considering $k = \frac{2\pi j}{N}$ and the range of values for such j , the solutions are

$$\begin{cases} k+q = \pi\ , \quad k \in K_1 \\ k'+q' = \pi\ , \quad k' \in K_1 \end{cases} \text{(if N even)} \quad \vee \quad \begin{cases} k = k' \\ q = q' \end{cases} \quad \vee \quad \begin{cases} k = q' \\ q = k' \end{cases} \tag{3.19}$$

where $k \in K_1$ means that $j_k \in [1, \frac{N}{2}-1] \cap \mathbb{N}$.

Second and third sub-systems

We go on with the second system,

$$\begin{cases} \omega(k)+\omega(q)=\omega(k')+\omega(q') \\ k+q=k'+q'+2\pi \end{cases} \implies \begin{cases} \cos(k)+\cos(q)=\cos(k')+\cos(q') \\ k+q=k'+q'+2\pi \end{cases}$$

$$\begin{cases} \cos(\frac{k+q}{2})\cos(\frac{k-q}{2})=\cos(\frac{k'+q'}{2})\cos(\frac{k'-q'}{2}) \\ k+q=k'+q'+2\pi \end{cases} \implies$$

$$\begin{cases} \cos(\frac{k+q}{2})\cos(\frac{k-q}{2})=\cos(\frac{k+q}{2}+\pi)\cos(\frac{k'-q'}{2}) \\ k+q=k'+q'+2\pi \end{cases} \implies$$

$$\begin{cases} \cos(\frac{k+q}{2})=0 \\ k+q=k'+q'+2\pi \end{cases} \vee \begin{cases} \cos(\frac{k-q}{2})=-\cos(\frac{k'-q'}{2}) \\ k+q=k'+q'+2\pi \end{cases}$$

Considering $k=\frac{2\pi j}{N}$ and the range of values for such j , the solutions are

$$\begin{cases} k+q=3\pi\ , & k\in K_2 \\ k'+q'=\pi\ , & k'\in K_1 \end{cases} \quad \text{(if N even)} \tag{3.20}$$

where $k\in K_2$ means that $j_k\in[\frac{N}{2},N]\cap\mathbb{N}$.

By symmetry, the third delta system

$$\begin{cases} \omega(k)+\omega(q)=\omega(k')+\omega(q') \\ k+q=k'+q'-2\pi \end{cases}$$

gives the solutions

$$\begin{cases} k+q=\pi\ , & k\in K_1 \\ k'+q'=3\pi\ , & k'\in K_2 \end{cases} \quad \text{(if N even)}\ . \tag{3.21}$$

Intersection of solutions

Let us tidy up all the results we found so far, labeling each set of solutions. Naturally, in our Hamiltonian these sets of solutions will be translated into a sum of deltas, in order to"pick" the right combinations of momenta from the sum over them. There is one thing to be careful about, though. Every time we sum two deltas pertaining to two different set of solutions, we over-count every combination that is part of both sets of solutions. Therefore, we need to actually see which solutions are in the intersection of more sets and take proper measure. We follow with labeling the solution sets and finding each

intersection — wherever there is no explicit restriction upon the values of the momenta, assume that it can take up any original value.

$$A = \begin{cases} k = k' \\ q = q' \end{cases} \qquad B = \begin{cases} k = q' \\ q = k' \end{cases} \qquad A \cap B = \begin{cases} k = k \\ k' = k \\ q = k \\ q' = k \end{cases}$$

$$C_1 = \begin{cases} k + q = \pi \ , \quad k \in K_1 \\ k' + q' = 1\pi \ , \quad k' \in K_1 \end{cases} \qquad C_2 = \begin{cases} k + q = 3\pi \ , \quad k \in K_2 \\ k' + q' = 3\pi \ , \quad k' \in K_2 \end{cases}$$

$$C_3 = \begin{cases} k + q = \pi \ , \quad k \in K_1 \\ k' + q' = 3\pi \ , \quad k' \in K_2 \end{cases} \qquad C_4 = \begin{cases} k + q = 3\pi \ , \quad k \in K_2 \\ k' + q' = \pi \ , \quad k' \in K_1 \end{cases}$$

$$A \cap C_1 = \begin{cases} k = k \\ k' = k \\ q = \pi - k \\ q' = \pi - k \end{cases} \ , \quad k \in K_1 \qquad A \cap C_2 = \begin{cases} k = k \\ k' = k \\ q = 3\pi - k \\ q' = 3\pi - k \end{cases} \ , \quad k \in K_2$$

$$B \cap C_1 = \begin{cases} k = k \\ k' = \pi - k \\ q = \pi - k \\ q' = k \end{cases} \ , \quad k \in K_1 \qquad A \cap C_2 = \begin{cases} k = k \\ k' = 3\pi - k \\ q = 3\pi - k \\ q' = k \end{cases} \ , \quad k \in K_2$$

$$A \cap C_3 = \varnothing \qquad A \cap C_4 = \varnothing \qquad B \cap C_3 = \varnothing \qquad B \cap C_4 = \varnothing$$

$$A \cap B \cap C_1 = \begin{cases} k = \frac{\pi}{2} \\ k' = \frac{\pi}{2} \\ q = \frac{\pi}{2} \\ q' = \frac{\pi}{2} \end{cases} \qquad A \cap B \cap C_2 = \begin{cases} k = \frac{3\pi}{2} \\ k' = \frac{3\pi}{2} \\ q = \frac{3\pi}{2} \\ q' = \frac{3\pi}{2} \end{cases}$$

Inserting the solutions

By calling each Hamiltonian term by the label associated to its solution set, let us insert the solutions into the Hamiltonian, in the form of a delta. We start with the main sets, noticing how the C_i sets are added together nicely once in the Hamiltonian.

$$A = -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\delta_{k,k'}\delta_{q,q'} = \\ = -\frac{g}{N}\sum_{kq} f_k^\dagger f_k f_q^\dagger f_q \quad , \tag{3.22}$$

$$B = -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\delta_{k,q'}\delta_{q,k'} = \\ = -\frac{g}{N}\sum_{kq} f_k^\dagger f_q f_q^\dagger f_k e^{-i(q-k)}. \tag{3.23}$$

We use the periodic properties of the fermions to be able to write $3\pi - k \to \pi - k$ every time. This makes it possible for us to merge together the following terms:

$$C_1 = -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\delta_{q,\pi-k}\delta_{q',\pi-k'} = \\ = -\frac{g}{N}\sum_{k,k'\in K_1} f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} \quad , \tag{3.24}$$

$$C_2 = -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\delta_{q,3\pi-k}\delta_{q',3\pi-k'} = \\ = -\frac{g}{N}\sum_{k,k'\in K_2} f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} \quad , \tag{3.25}$$

$$C_3 = -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\delta_{q,\pi-k}\delta_{q',3\pi-k'} = \\ = -\frac{g}{N}\sum_{\substack{k\in K_1\\ k'\in K_2}} f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} \quad , \tag{3.26}$$

$$C_4 = -\frac{g}{N}\sum_{kk'}\sum_{qq'} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\delta_{q,\pi-k}\delta_{q',\pi-k'} = \\ = -\frac{g}{N}\sum_{\substack{k\in K_2\\ k'\in K_1}} f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} \quad . \tag{3.27}$$

As we said, they are added together like so:

$$C = C_1 + C_2 + C_3 + C_4 = \\ = -\frac{g}{N}\Big[\sum_{k,k'\in K_1} + \sum_{k,k'\in K_2} + \sum_{\substack{k\in K_1\\ k'\in K_2}} + \sum_{\substack{k\in K_2\\ k'\in K_1}}\Big] f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} = \\ = -\frac{g}{N}\Big[\sum_{k\in K_1}\sum_{k'} + \sum_{k\in K_2}\sum_{k'}\Big] f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} = \\ = -\frac{g}{N}\Big[\sum_{k}\sum_{k'}\Big] f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} = -\frac{g}{N}\sum_{kk'} f_k^\dagger f_{k'} f_{\pi-k}^\dagger f_{\pi-k'} e^{i(k-k')} \quad . \tag{3.28}$$

We can now deal with the intersection terms:

$$A \cap B = -\frac{g}{N}\sum_k f_k^\dagger f_k f_k^\dagger f_k = -\frac{g}{N}\sum_k f_k^\dagger f_k \quad , \tag{3.29}$$

$$A \cap C = A \cap C_1 + A \cap C_2 = \\ = -\frac{g}{N}\Big[\sum_{k\in K_1} f_k^\dagger f_k f_{\pi-k}^\dagger f_{\pi-k} + \sum_{k\in K_2} f_k^\dagger f_k f_{3\pi-k}^\dagger f_{3\pi-k}\Big] = \\ = -\frac{g}{N}\sum_k f_k^\dagger f_k f_{\pi-k}^\dagger f_{\pi-k} \quad , \tag{3.30}$$

$$B \cap C = B \cap C_1 + B \cap C_2 = \\ = +\frac{g}{N}\Big[\sum_{k\in K_1} f_k^\dagger f_{\pi-k} f_{\pi-k}^\dagger f_k e^{2ik} + \sum_{k\in K_2} f_k^\dagger f_{3\pi-k} f_{3\pi-k}^\dagger f_k e^{2ik}\Big] = \\ = +\frac{g}{N}\sum_k f_k^\dagger f_{\pi-k} f_{\pi-k}^\dagger f_k e^{2ik} \quad , \tag{3.31}$$

$$A \cap B \cap C = A \cap B \cap C_1 + A \cap B \cap C_2 =$$

$$= -\frac{g}{N} f^\dagger_{\frac{\pi}{2}} f_{\frac{\pi}{2}} - \frac{g}{N} f^\dagger_{\frac{3\pi}{2}} f_{\frac{3\pi}{2}} \quad . \tag{3.32}$$

With the many properties we have available in the fermion world, we really have freedom in terms of how we want our Hamiltonian pieces to look. We try our best to rewrite some of them in such a way to later have terms cancel, while still retaining a good view of how each term will behave. We rewrite B in the following way:

$$\begin{aligned} B &= -\frac{g}{N} \sum_k f^\dagger_k f_k + \frac{g}{N} \sum_{kq} f^\dagger_k f_k f^\dagger_q f_q e^{-i(q-k)} = \\ &= -\frac{g}{N} \sum_k f^\dagger_k f_k + \frac{g}{N} \Big| \sum_k f^\dagger_k f_k e^{ik} \Big|^2 \quad ; \end{aligned} \tag{3.33}$$

followed by $B \cap C$,

$$\begin{aligned} B \cap C &= +\frac{g}{N} \sum_k f^\dagger_k f_{\pi-k} f^\dagger_{\pi-k} f_k e^{2ik} = \\ &= ... = +\frac{g}{N} \sum_k f^\dagger_k f_k e^{2ik} - \frac{g}{N} \sum_k f^\dagger_k f_k f^\dagger_{\pi-k} f_{\pi-k} e^{2ik} \\ &- \frac{g}{N} f^\dagger_{\frac{\pi}{2}} f_{\frac{\pi}{2}} - \frac{g}{N} f^\dagger_{\frac{3\pi}{2}} f_{\frac{3\pi}{2}} \quad , \end{aligned} \tag{3.34}$$

and C itself:

$$\begin{aligned} C &= \frac{g}{N} \sum_{kk'} f^\dagger_k f_{k'} f^\dagger_{\pi-k} f_{\pi-k'} e^{i(k-k')} = \\ &= ... = \frac{g}{N} \sum_k f^\dagger_k f_k e^{2ik} - \frac{g}{N} \Big| \sum_k f^\dagger_k f^\dagger_{\pi-k} e^{ik} \Big|^2 \quad . \end{aligned} \tag{3.35}$$

For the sake of clarity, we list all of the Hamiltonian terms here, one by one, in the shape we last chose:

$$A = -\frac{g}{N} \sum_{kq} f^\dagger_k f_k f^\dagger_q f_q$$

$$B = -\frac{g}{N} \sum_k f^\dagger_k f_k + \frac{g}{N} \Big| \sum_k f^\dagger_k f_k e^{ik} \Big|^2$$

$$C = \frac{g}{N} \sum_k f^\dagger_k f_k e^{2ik} - \frac{g}{N} \Big| \sum_k f^\dagger_k f^\dagger_{\pi-k} e^{ik} \Big|^2$$

$$A \cap B = -\frac{g}{N} \sum_k f^\dagger_k f_k$$

$$A \cap C = -\frac{g}{N} \sum_k f^\dagger_k f_k f^\dagger_{\pi-k} f_{\pi-k}$$

$$B \cap C = +\frac{g}{N}\sum_k f_k^\dagger f_k e^{2ik} - \frac{g}{N}\sum_k f_k^\dagger f_k f_{\pi-k}^\dagger f_{\pi-k} e^{2ik} - \frac{g}{N} f_{\frac{\pi}{2}}^\dagger f_{\frac{\pi}{2}} - \frac{g}{N} f_{\frac{3\pi}{2}}^\dagger f_{\frac{3\pi}{2}}$$
$$A \cap B \cap C = -\frac{g}{N} f_{\frac{\pi}{2}}^\dagger f_{\frac{\pi}{2}} - \frac{g}{N} f_{\frac{3\pi}{2}}^\dagger f_{\frac{3\pi}{2}}$$

3.2.2 First order normal form

We are now ready to put together all that we have found, to write a version of the first order normal form of the perturbation in terms that are, more or less, manifest in their action. The reasoning behind the first line of the following calculation comes naturally from trying to *uniformly cover* three sets which have all nonzero intersection. We follow with the whole explicit computation, cancelling all the terms that do:

$$\begin{aligned}
\hat{S} &= A + B + C - A\cap B - A\cap C - B\cap C + A\cap B\cap C = \\
&= -\frac{g}{N}\sum_{kq} f_k^\dagger f_k f_q^\dagger f_q - \frac{g}{N}\sum_k f_k^\dagger f_k + \frac{g}{N}\Big|\sum_k f_k^\dagger f_k e^{ik}\Big|^2 + \frac{g}{N}\sum_k f_k^\dagger f_k e^{2ik} \\
&- \frac{g}{N}\Big|\sum_k f_k^\dagger f_{\pi-k}^\dagger e^{ik}\Big|^2 + \frac{g}{N}\sum_k f_k^\dagger f_k + \frac{g}{N}\sum_k f_k^\dagger f_k f_{\pi-k}^\dagger f_{\pi-k} - \frac{g}{N}\sum_k f_k^\dagger f_k e^{2ik} \\
&+ \frac{g}{N}\sum_k f_k^\dagger f_k f_{\pi-k}^\dagger f_{\pi-k} e^{2ik} + \frac{g}{N} f_{\frac{\pi}{2}}^\dagger f_{\frac{\pi}{2}} + \frac{g}{N} f_{\frac{3\pi}{2}}^\dagger f_{\frac{3\pi}{2}} - \frac{g}{N} f_{\frac{\pi}{2}}^\dagger f_{\frac{\pi}{2}} - \frac{g}{N} f_{\frac{3\pi}{2}}^\dagger f_{\frac{3\pi}{2}} = \\
&= -\frac{g}{N}\sum_{kq} f_k^\dagger f_k f_q^\dagger f_q + \frac{g}{N}\Big|\sum_k f_k^\dagger f_k e^{ik}\Big|^2 - \frac{g}{N}\Big|\sum_k f_k^\dagger f_{\pi-k}^\dagger e^{ik}\Big|^2 \\
&+ \frac{g}{N}\sum_k f_k^\dagger f_k f_{\pi-k}^\dagger f_{\pi-k} + \frac{g}{N}\sum_k f_k^\dagger f_k f_{\pi-k}^\dagger f_{\pi-k} e^{2ik} = \\
&= -\frac{g}{N}\Big(\sum_k f_k^\dagger f_k\Big)^2 + \frac{g}{N}\Big|\sum_k f_k^\dagger f_k e^{ik}\Big|^2 - \frac{g}{N}\Big|\sum_k f_k^\dagger f_{\pi-k}^\dagger e^{ik}\Big|^2 \\
&+ \frac{g}{N}\sum_k \Big|f_k^\dagger f_{\pi-k}^\dagger\Big|^2\big(1+\cos(2k)\big) = \\
&= -\frac{g}{N}\sum_{kq} f_k^\dagger f_k f_q^\dagger f_q\big(1-\cos(k-q)\big) - \frac{g}{N}\sum_{kq} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q) \\
&+ \frac{g}{N}\sum_k f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k\big(1+\cos(2k)\big) \quad .
\end{aligned}$$

Notice that $\hat{S}$ here is manifestly hermitian. Also, recall that each term in which $\pi - k$ appears comes from the C sets of solutions, which are present only in the case of N even. For this reason, let us neatly restate our result, highlighting the fact that a part of it appears only whenever N is divisible

by 2:

$$\begin{aligned}\hat{S} = &-\frac{g}{N}\sum_{kq} f_k^\dagger f_k f_q^\dagger f_q \big(1-\cos(k-q)\big) \\ &-\frac{g}{N}\sum_{kq} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q)\delta_{N\%2} \\ &+\frac{g}{N}\sum_{k} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k \big(1+\cos(2k)\big)\delta_{N\%2} \quad . \end{aligned} \tag{3.36}$$

To simplify later calculations, we also give a label to each of the three terms here present:

$$\hat{S} = S_1 + S_{N\%2} = S_1 + S_A + S_B \quad , \tag{3.37}$$

where

$$S_1 = -\frac{g}{N}\sum_{kq} f_k^\dagger f_k f_q^\dagger f_q \big(1-\cos(k-q)\big) \quad ; \tag{3.38}$$

$$S_A = -\frac{g}{N}\sum_{kq} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q)\delta_{N\%2} \quad ; \tag{3.39}$$

$$S_B = \frac{g}{N}\sum_{k} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k \big(1+\cos(2k)\big)\delta_{N\%2} \quad . \tag{3.40}$$

Checking consistency

From Hamiltonian perturbation theory explored in Appendix A, we should have that $[\hat{S}, \hat{h}] = 0$, if we did everything correctly. Let us now check just that, as it is not really obvious in the version of S we just obtained. We will use the splitting in (3.37) to break the calculations into three easier chunks. The first, naturally, is zero, since every addend in (3.38) is a function of some number operator, exactly like h, and we know that number operators always commute. Therefore,

$$[S_1, \hat{h}] = 0 \quad . \tag{3.41}$$

The second term is definitely trickier. Let us first see how (3.39) commutes with a generic number operator:

$$\begin{aligned}[S_A, n_p] = &-\frac{g}{N}\sum_{kq}[f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q, n_p]\cos(k+q) = \\ = &-\frac{g}{N}\sum_{kq}\big([f_k^\dagger, n_p] f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q) + f_k^\dagger [f_{\pi-k}^\dagger, n_p] f_{\pi-q} f_q \cos(k+q) + \end{aligned}$$

$$
\begin{aligned}
&+ f_k^\dagger f_{\pi-k}^\dagger [f_{\pi-q}, n_p] f_q \cos(k+q) + f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} [f_q, n_p] \cos(k+q)\big) = \\
=&-\frac{g}{N}\sum_{kq}\big(- \delta_{k,p} f_p^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q) - f_k^\dagger \delta_{\pi-k,p} f_p^\dagger f_{\pi-q} f_q \cos(k+q) + \\
&+ f_k^\dagger f_{\pi-k}^\dagger \delta_{\pi-q,p} f_p f_q \cos(k+q) + f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} \delta_{q,p} f_p \cos(k+q)\big) = \\
=&\frac{g}{N}\sum_{q}\big(f_p^\dagger f_{\pi-p}^\dagger f_{\pi-q} f_q \cos(p+q) + f_p^\dagger f_{\pi-p}^\dagger f_{\pi-q} f_q \cos(q-p)\big) + \\
&-\frac{g}{N}\sum_{k}\big(f_k^\dagger f_{\pi-k}^\dagger f_{\pi-p} f_p \cos(k-p) + f_k^\dagger f_{\pi-k}^\dagger f_{\pi-p} f_p \cos(k+p)\big) = \\
=&\frac{g}{N}\sum_{q}\cos(p+q)\big(- f_p^\dagger f_{\pi-p}^\dagger f_{\pi-q} f_q + f_q^\dagger f_{\pi-q}^\dagger f_{\pi-p} f_p\big) + \\
&\frac{g}{N}\sum_{q}\cos(p-q)\big(- f_p^\dagger f_{\pi-p}^\dagger f_{\pi-q} f_q + f_q^\dagger f_{\pi-q}^\dagger f_{\pi-p} f_p\big) = \\
=&-\frac{g}{N}\sum_{q}\big[f_p^\dagger f_{\pi-p}^\dagger f_{\pi-q} f_q - f_q^\dagger f_{\pi-q}^\dagger f_{\pi-p} f_p\big] 2\cos(p)\cos(q) \quad . \qquad (3.42)
\end{aligned}
$$

We will leave it like this for now, and also calculate the commutator between (3.40) and the number operator:

$$
\begin{aligned}
[S_B, n_p] =&\frac{g}{N}\sum_{k}[f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k, n_p]\big(1+\cos(2k)\big) = \\
=&\frac{g}{N}\sum_{k}\Big[- \delta_{k,p} f_p^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k\big(1+\cos(2k)\big) - \delta_{\pi-k,p} f_k^\dagger f_p^\dagger f_{\pi-k} f_k\big(1+\cos(2k)\big) + \\
&+ \delta_{\pi-k,p} f_k^\dagger f_{\pi-k}^\dagger f_p f_k\big(1+\cos(2k)\big) + \delta_{k,p} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_p\big(1+\cos(2k)\big)\Big] = \\
=&\frac{g}{N}\sum_{k}\Big[- f_p^\dagger f_{\pi-p}^\dagger f_{\pi-p} f_p\big(1+\cos(2p)\big) - f_p^\dagger f_{\pi-p}^\dagger f_{\pi-p} f_p\big(1+\cos(2p)\big) + \\
&+ f_p^\dagger f_{\pi-p}^\dagger f_{\pi-p} f_p\big(1+\cos(2p)\big) + f_p^\dagger f_{\pi-p}^\dagger f_{\pi-p} f_p\big(1+\cos(2p)\big)\Big] = \\
=&\ 0 \quad . \qquad (3.43)
\end{aligned}
$$

Then, combining (3.42) and (3.43), we find that

$$
\begin{aligned}
[S_{N\%2}, \hat{h}] =&\sum_{k,q}\omega(k)[S_B, n_k] = \\
=&-\frac{g^2}{N}\sum_{k,q}\big(1-\cos(k)\big)\big[f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q - f_q^\dagger f_{\pi-q}^\dagger f_{\pi-k} f_k\big] 2\cos(k)\cos(q) \quad .
\end{aligned}
$$

By exchanging $k \leftrightarrow q$ in one half of this term, the contribution of 1 into $\omega(k)$ disappears, and we are left with

$$[S_{N\%2}, \hat{h}] = -\frac{g^2}{N} \sum_{k,q} \; f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q 2\cos(k)\cos(q)[\cos(k) - \cos(q)] \quad ,$$

which, if we send $k \to \pi - k$ and $q \to \pi - q$, is sent into its exact opposite. This is always allowed, and we can change the indexes in half of it to prove that it is zero. We finally found, as we anticipated, that

$$[\hat{S}, \hat{h}] = 0 \quad . \tag{3.44}$$

3.2.3 On the dynamics of the system

Let us now briefly discuss the implications of the perturbative approach to the dynamic of the system.

Notice that while we were proving that $[\hat{S}, \hat{h}] = 0$ just in the previous paragraph, we also proved that $[S, \sum_k n_k] = 0$, which tells us that the conservation of the whole number of fermion modes in the system — that is related to the conservation of total z component in the spin chain — is still valid in the new system described by the perturbed Hamiltonian $H = h + S$. Basically, this tells us that our process has not disrupted a fundamental property of the system, even though we truncated part of its original Hamiltonian.

Additionally, for the N odd case we found a result which is in complete analogy with what usually happens in the classical system setting of traditional Hamiltonian perturbation theory dealing with the symplectic Poisson tensor and with action-angle variables. When the unperturbed classical Hamiltonian is a function of the sole action variables, it might happen — and this greatly simplifies the problem — that an eventual dependence on the angle variable disappears from the perturbation once it is put in truncated normal form (usually at first order). In general, this is a recurrent result in some applications of classical Hamiltonian perturbation theory, and our quantum case is no exception: the normal form of the perturbation in the case of odd number of particles (or sites, or spins) is

$$S_1 = -\frac{g}{N} \sum_{kq} f_k^\dagger f_k f_q^\dagger f_q \big(1 - \cos(k - q)\big) \quad , \tag{3.45}$$

and clearly depends only on the number operators n_k, which plays the role of the action variable, the absolute value, etc. , and not on single creation or annihilation operators. Just as in the classical case, where the dependence on the action variable only implies a new set of first integrals with respect

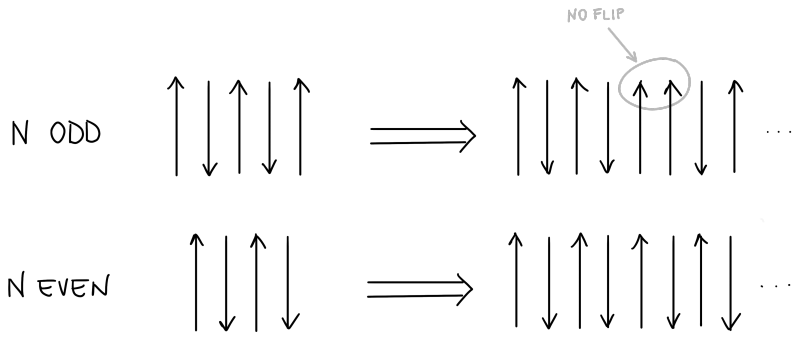
to the original system, here too we find that the dynamic of the system is quite trivial, and that we have a whole new collection of conserved quantities. In fact, the Hamiltonian $\hat{H} = \hat{h} + S_1$ commutes with every single number operator (recall that number operators commute with each other), which means that every one of them is separately conserved; this does not happen in the original system. Going back to the spin configuration, it would mean that each z-component of the z is conserved and does not change.

In the N even case, though, this simplification does not happen: the term

$$S_{N\%2} = -\frac{g}{N}\left[\sum_{kq} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q) - \sum_k f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k \big(1+\cos(2k)\big)\right] \tag{3.46}$$

does not commute with single number operators, but only with certain combinations of them. This corresponds, in a classical case, to having a perturbation in normal form that retains some dependence on the angle variables, particularly in a way that implies a change in time of the actions. In terms of spins, this means that when N is odd, the z-component of each spin can vary in time, even in the first order perturbed system.

The fundamental reason for this non-negligible difference between the even and odd cases is still somewhat mysterious to us. We can certainly tell that there is a major difference between the two, both in terms of possible momenta and of possible configurations. In the Fourier space framework, we notice that in the case of even N there is the appearance of a new possible momentum $k = \pi$, which adds a completely new degree of degeneracy and makes it possible to have more solutions for the delta system, resulting in the additional terms. The new type degeneracy found is further discussed in 3.3.3. In the spin configuration setting, these additional terms in the normal form of the perturbation seem to be connected to a particular type of configuration that is possible only with an even number of sites: the one where all the neighbors of any spin are flipped, which would be the ground state for a complete anti-ferromagnetic Heisenberg spin chain with no external field. Bear in mind that while it is always possible, naturally, to have a state with spin up-down-up-down, $|\uparrow\downarrow\uparrow\downarrow\uparrow ...\rangle$, we are working with periodic conditions. Because of that, when we spread out the N sites into an infinitely long chain, we see right away that the complete anti-ferromagnetic ground state is possible only when N is even, as illustrated in the following sketch:



Actually, this is just a particular case of configuration possible only on even sites. More generally, any configuration of repeating intervals (up or down) of the same length is impossible in the case of odd N. At a physical level, if we actually visualize the circular version of the chain, we see that what is missing in the odd case is the configurations which are conserved under spin flip, i.e. the "balanced" configurations where, naturally, the total z-component of the chain is zero. This could have been seen right away, since the total z-component can only assume values in $[-\frac{N}{2}, \frac{N}{2}]$ distributed at unit intervals, so that the value $S^z = 0$ is absent when N is odd.

Therefore, the two cases really appear to be structurally different. What remains to understand, besides a deeper understanding of this feature, is why techniques like the Bethe ansatz seem to not mention this at all; meaning, if this is purely a result of the perturbation theory — although it does not look like it, given how there is an intuitive difference between the two cases — or if this discrepancy is actually present in the traditional method, but hidden somewhere.

3.3 Eigenvalue corrections

Now that we found the normal form of the perturbation, since we are working in the quantum regime, it is natural to be interested in the energy levels of the system. Finding the eigenvalues of the Hamiltonian is important in order to better understand the dynamic of the system and the role of each state, as well as evaluate the possible advantages and disadvantages of our approach by comparing the levels found with known results, or with the levels of the unperturbed system. Naturally, the eigenvalues of the energy are also a starting point for any study involving the statistical mechanics of the system, as they are needed to compute the partition function.

What we propose to do is to therefore give our approximate version of the eigenvalues of the Hamiltonian. Alas, we do have to make the distinction between even and odd sites/modes — a distinction which is absent in other kinds of approaches — and which will make the computation of the eigenvalues straightforward in one case, but a terrible pain in the other.

A generally good thing about finding the normal form of the perturbation to some order is that it may greatly simplify later computations by means of the usual time-independent perturbation theory — the same old theory for eigenvalue corrections, present in any good book on matter physics. This is especially true whenever the unperturbed system has a non-degenerate spectrum. Because of the commutation property between S and h, we have, for general eigenstates of h $|k_{1,2}\rangle$,

$$\langle k_1| \left[\hat{S}, \hat{h}\right] |k_2\rangle = \langle k_1| \left[\hat{S}E(k_2) - E(k_1)\hat{S}\right] |k_2\rangle \stackrel{!}{=} 0 \quad , \tag{3.47}$$

which implies, if $E(k_2) \neq E(k_1)$,

$$\langle k_1| S |k_2\rangle = 0 \quad . \tag{3.48}$$

Unfortunately, this is not the case for our $\hat{h}$. The degeneracy does exist, smaller in the case of N odd and greater and more complicated in the case of N even. Nevertheless, the incredible property of $\hat{S}$ will prove to lighten part of the cumbersome calculations in our case as well.

3.3.1 Fermionic Fock states

Before we continue, let us freshen up on the state basis upon which our fermionic operators act. Obviously, since we are working with creation, annihilation and number operators, the proper Hilbert space is a Fock space $\mathcal{F}$ of the form

$$\mathcal{F} = \bigoplus_{M=0}^{N} \mathcal{H}_M \quad , \tag{3.49}$$

where M indicates the number of fermions present in the state. The $\mathcal{H}_0$ block here spans a unique state with $M = 0$, the vacuum state $|0\rangle := |\text{all } \mathrm{n}_\alpha \text{ are zero}\rangle$. The $\mathcal{H}_1$ block spans those states with $M = 1$, where only a single $n_\alpha = 1$ and all the other n_β with $\beta \neq \alpha$ are zero. Similar to the very well known bosonic case, we may identify the states with the notation below

$$|\alpha\rangle := |n_\alpha = 1; \text{any other } n = 0\rangle = f_\alpha^\dagger |0\rangle \quad , \tag{3.50}$$

therefore identifying the $\mathcal{H}_1$ block of the Fock space with a single particle Hilbert space. The $\mathcal{H}_2$ block spans those states with $M = 2$, where only two indexes α and β are such that $n_\alpha = 1$, $n_\beta = 1$, while all the others are zero. These states ($|n_\alpha = 1; n_\beta = 1; \text{any other } n = 0\rangle$) will be written as $|\alpha, \beta\rangle$. One has to be careful with such states now that we have more than one particle, as the commutation relations of the fermionic operators (and thus the anti-symmetric properties — under exchange of particles — of any two-body fermion wavefunction) has the following implication:

$$|\alpha, \beta\rangle := f^\dagger_\beta f^\dagger_\alpha |0\rangle = -f^\dagger_\alpha f^\dagger_\beta |0\rangle = -|\beta, \alpha\rangle \quad . \tag{3.51}$$

This means that, if out of N spots we have $M = 2$ possible particles α and β present, we have the following regarding the orthonormality of the states in $\mathcal{H}_2$:

$$\langle \alpha_1, \beta_1 | \alpha_2, \beta_2 \rangle = \delta_{\alpha_1,\alpha_2}\delta_{\beta_1,\beta_2} - \delta_{\alpha_1,\beta_2}\delta_{\beta_1,\alpha_2} \quad . \tag{3.52}$$

We can continue the discussion analogously for each $\mathcal{H}_M$, where we only have to be careful that, in all cases, the order of the particles does not matter physically but affects the overall sign of the state according to the sign of the permutation from the original.

Let us make just one more point about the use of these states, now in relation with the annihilation operator. Since we have defined the states in (3.51) conforming to the action of consecutive creation operators upon the vacuum state, we require a better understanding of the annihilation operator on a given state labeled by $|\alpha_{i_1}, ..., \alpha_{i_M}\rangle$, where the indexes i_a are ordered but not necessarily the whole set $[1, M] \cap \mathbb{N}$.

Then, the annihilation of the i_b-th fermion would look like:

$$f_{\alpha_{i_b}} |\alpha_{i_1}, ..., \alpha_{i_b}, ..., \alpha_{i_M}\rangle = (\delta_{n_{\alpha_{i_b}},1})(-1)^{\sum_{a>i_b} n_a} |\alpha_{i_1}, ..., \sim \alpha_{i_b}, ..., \alpha_{i_M}\rangle \quad , \tag{3.53}$$

where $\sim \alpha_{i_b}$ means that the once present fermion in that spot is now gone. The factor $(-1)^{\sum_{a>i_b} n_a}$ is called the *Jordan-Wigner string*. With this construction, we have that each $\mathcal{H}_M$ with $M \geq 2$ is a Hilbert space of M identical fermions with well defined properties and responses under fermionic operators.

Eigenstates of $\hat{h}$

The conservation of the total z-component of the spin in our system is, as we know, reversed into the conservation of total number of fermions, or fermion

modes. This means that whenever the system Hamiltonian acts upon a state inside $\mathcal{H}_M$, the resulting state will again be part of $\mathcal{H}_M$. In the specific case of the unperturbed Hamiltonian h, we have even more restrictive conditions: since h is solely a function of the n_k's, which are all conserved as they commute with each other, we have that any fermionic Fock state here mentioned is actually an eigenstate of h.

For the sake of clarity, let us write the explicit convention for the states that will be used from now on in the discussion. A state with M fermion modes present with wave numbers $k_1, ..., k_M$ will be denoted as

$$|k_1, k_2, ..., k_M\rangle \quad \text{such that} \quad |k, q\rangle = -\,|q, k\rangle \quad . \tag{3.54}$$

Then, the operators act as such:

$$n_q\,|k_1, k_2, ..., k_M\rangle = \sum_j \delta_{q,k_j}\,|k_1, k_2, ..., k_M\rangle \ , \tag{3.55}$$

$$f_q\,|k_1, k_2, ..., k_M\rangle = \sum_j \delta_{q,k_j}(-1)^{M-j}\,|k_1, k_2, ..., (\sim k_j), ..., k_M\rangle \ , \tag{3.56}$$

$$f_q^\dagger\,|k_1, k_2, ..., k_M\rangle = \Big[\prod_j (1-\delta_{q,k_j})\Big]\,|k_1, k_2, ..., k_M, q\rangle \ , \tag{3.57}$$

where the factor $(-1)^{M-j}$ serves the purpose of the Jordan-Wigner string.

The unperturbed Hamiltonian $\hat{h}$ applied to these state gives then the unperturbed spectrum:

$$\begin{aligned} h\,|k_1, k_2, ..., k_M\rangle &= g\sum_{j=1}^{M}(1-\cos k_j)\,|k_1, k_2, ..., k_M\rangle = \\ &= g\Big[M - \sum_{j=1}^{M}\cos k_j\Big]\,|k_1, k_2, ..., k_M\rangle \quad . \end{aligned} \tag{3.58}$$

3.3.2 Odd number of modes N

The case of odd number of total possible modes N proves to be the simplest one to study, since the portion of S that survives in this case has the same eigenvectors of the unperturbed system. Indeed, as one can see right away, the first order normal form here is written only in terms of single mode number operators, just like $\hat{h}$:

$$S_1 = -\frac{g}{N}\sum_{kq} n_k n_q\big(1-\cos(k-q)\big) \quad . \tag{3.59}$$

Let us mention that the spectrum of the unperturbed system is degenerate in this case for any substitution $k \longrightarrow 2\pi - k$. For simplicity let us just call any two such states $|k, ...\rangle$ and $|2\pi - k, ...\rangle$. Such degeneration is not canceled by the perturbation, as S_1 commutes with an operator which actually would split the degeneration ($\mathcal{K} = \sum_k k n_k$) and therefore

$$\langle k, ...| \, [S_1, \mathcal{K}] \, |2\pi - k, ...\rangle = \Big(2\pi - k + (...) - k - (...)\Big) \, \langle k, ...| \, S_1 \, |2\pi - k, ...\rangle = \\ \stackrel{!}{=} 0 \quad , \tag{3.60}$$

which implies

$$\langle k, ...| \, S_1 \, |2\pi - k, ...\rangle = 0 \quad . \tag{3.61}$$

Thus, the eigenvalue corrections are simply given by the equation

$$S_1 \, |k_1, ..., k_M\rangle = E_2(k_1, ..., k_M) \, |k_1, ..., k_M\rangle \quad . \tag{3.62}$$

The calculation follows this way:

$$\begin{aligned} E_2(k_1, ..., k_M) \, |k_1, ..., k_M\rangle &= -\frac{g}{N} \sum_{kq} n_k n_q \big(1 - \cos(k - q)\big) \, |k_1, k_2, ..., k_M\rangle = \\ &= -\frac{g}{N} \sum_{k} n_k \sum_{j} \big(1 - \cos(k - k_j)\big) \, |k_1, k_2, ..., k_M\rangle = \\ &= -\frac{g}{N} \sum_{i} \sum_{j} \big(1 - \cos(k_i - k_j)\big) \, |k_1, k_2, ..., k_M\rangle = \\ &= -\frac{g}{N} \Big[M^2 - M - 2 \sum_{i} \sum_{j,j>i} \cos(k_i - k_j)\Big] \, |k_1, k_2, ..., k_M\rangle \quad . \end{aligned}$$

We lastly get the correction to the spectrum in the case of odd N, for a generic number $M \leq N$ of modes present:

$$E_2(k_1, ..., k_M) = -\frac{g}{N} \Big[M^2 - M - 2 \sum_{i} \sum_{j,j>i} \cos(k_i - k_j)\Big] \quad . \tag{3.63}$$

3.3.3 Even number of modes N

The case for N even is much more problematic. This is due to the fact that we can now witness the appearance of a mode which was absent before, the mode with wave number $q = \pi$, and of the pairings $\{k, \pi - k\}$. Such pairs were of course not present earlier among the discrete set of possible wave numbers. This is crucial because it introduces a new degree of degeneracy in

the unperturbed spectrum. This degeneracy proves to be quite a challenge to handle, for it groups any state containing a pair $\{k, \pi - k\}$ with all the other states in which some $\{q, \pi - q\}$ is present instead of the k pair, while everything else remains untouched. We can see this very quickly in the following:

$$\hat{h}\,|k, \pi - k, \text{something}\rangle = \Big[E(\text{something}) + g(2 - \cos(k) - \cos(\pi - k)\Big] \\ |k, \pi - k, \text{something}\rangle = \tag{3.64}$$

$$= \Big[E(\text{something}) + 2g\Big]\,|k, \pi - k, \text{something}\rangle \quad . \tag{3.65}$$

It is clear then that q can have any possible value, excluding whatever modes are included it the "something", and still be part of the same degenerate subspace of the original state. This means, especially in the case of low spin-flip (i.e. few modes M) and large N, that we have a really huge degree of degeneracy, to be added to the $k \longrightarrow 2\pi - k$ degeneracy discussed in the last section.

To recap, this is what we can say about what kind of states are degenerate to $|k, \pi - k, \text{something}\rangle$:

$$|k, \pi - k, \text{something}\rangle \longleftrightarrow \begin{cases} |2\pi - k, \pi - k, \text{something}\rangle \\ |k, \pi + k, \text{something}\rangle \\ |q, \pi - q, \text{something}\rangle \\ |2\pi - q, \pi - q, \text{something}\rangle \\ |q, \pi + q, \text{something}\rangle \end{cases} , \tag{3.66}$$

where each and every one of these states is also degenerate to the version of themselves in which any of the elements in "something" are sent to their "$2\pi-$" version. Thus, the dimension of each degenerate subspace varies greatly upon the content of "something" and its relation with the pair/s. You can see how very confusing it all looks already.

For each degenerate subspace of dimension d, we can write the first order correction eigenvalue system according to the theory of time-independent

perturbations in case of degenerate spectrum:

$$\begin{cases} \mathcal{S}_{11}\alpha_1 + \mathcal{S}_{12}\alpha_2 + ... + \mathcal{S}_{1d}\alpha_d = E^{(1)}\alpha_1 \\ \mathcal{S}_{21}\alpha_1 + \mathcal{S}_{22}\alpha_2 + ... + \mathcal{S}_{2d}\alpha_d = E^{(1)}\alpha_2 \\ . \\ . \\ . \\ \mathcal{S}_{d1}\alpha_1 + \mathcal{S}_{d2}\alpha_2 + ... + \mathcal{S}_{dd}\alpha_d = E^{(1)}\alpha_d \end{cases} , \tag{3.67}$$

where $\mathcal{S}_{ij} = \langle\psi_i| S |\psi_j\rangle$ and the $|\psi_i\rangle$ are the eigenstates in the degenerate subspace in consideration.

Sadly, if in the odd case we found that the matrix $\mathcal{S}$ is diagonal, which made computations easy even though the degeneracy was not removed, in this case it is not. Recall that $\hat{S}$ is made up of two parts, one which is the only one that survives in the odd case, see (3.59), and one which acts on a state only when a pair $\{k, \pi - k\}$ is present:

$$\hat{S} = S_1 + S_{N\%2} \quad . \tag{3.68}$$

Whenever we have a degeneracy of the first type, i.e. $k \to 2\pi - k$, that makes all the pairs $(k, \pi - k)$ disappear, the only contribution that matters is once again S_1, and the relevant properties are the same as in the odd case: a diagonal block of equal entries in $\mathcal{S}$. In the case of pair degeneracy, we see that for different states with pairs associated respectively to k and q, the matrix element is now different from zero, as the section $S_{N\%2}$ couples exactly such states, with a non zero contribution. The S_1 part of the Hamiltonian is here absent, since any state is its eigenstate, and thus there is no coupling. The diagonal entries for the states with at least one pair are interesting: the S_1 contribution is not zero, while the $S_{N\%2}$ contribution is, although it might not be obvious at a first glance. We will address the proof of this at the end of this section.

As one can see, there is not much hope to formally solve a system like this. It varies greatly with the kind and number of modes present in the states.

Proof of zero contribution for $S_{N\%2}$ in diagonal pair terms of $\mathcal{S}$

Let us first split $S_{N\%2}$ into two more parts:

$$\begin{aligned} S_{N\%2} &= S_A + S_B \\ &= -\frac{g}{N}\sum_{kq} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q) + \frac{g}{N}\sum_k f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k \big(1 + \cos(2k)\big) \quad , \end{aligned}$$

where

$$S_A = -\frac{g}{N}\sum_{kq} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q) \quad , \tag{3.69}$$

$$S_B = +\frac{g}{N}\sum_{k} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k \big(1+\cos(2k)\big) \quad . \tag{3.70}$$

Then, the contribution of the $S_{N\%2}$ part of $\hat{S}$ into the diagonal entries of $\mathcal{S}$, associated with states in which a pair $\{k, \pi-k\}$ is present, is given by

$$\begin{aligned}\langle \bar{k}, \pi-\bar{k}|\, S_{N\%2}\, |\bar{k}, \pi-\bar{k}\rangle = \langle \bar{k}, \pi-\bar{k}|\, S_A\, |\bar{k}, \pi-\bar{k}\rangle + \\ + \langle \bar{k}, \pi-\bar{k}|\, S_B\, |\bar{k}, \pi-\bar{k}\rangle \quad .\end{aligned} \tag{3.71}$$

Let us handle the two additions separately, remaining well aware of the unintuitive properties of fermionic Fock states:

$$\begin{aligned}\langle \bar{k}, \pi-\bar{k}|\, S_A\, |\bar{k}, \pi-\bar{k}\rangle &= -\frac{g}{N}\,\langle \bar{k}, \pi-\bar{k}| \sum_{kq} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-q} f_q \cos(k+q)\, |\bar{k}, \pi-\bar{k}\rangle \\ &= -\frac{g}{N}\,\langle \bar{k}, \pi-\bar{k}| \sum_{kq} f_k^\dagger f_{\pi-k}^\dagger \cos(k+q)(\delta_{q,\pi-\bar{k}} - \delta_{q,\bar{k}})\, |0\rangle \\ &= +\frac{g}{N}\sum_k [\cos(k+\bar{k}) + \cos(k-\bar{k})]\,\langle \bar{k}, \pi-\bar{k}|\, |\pi-k, k\rangle \\ &= +\frac{g}{N}\sum_k [\cos(k+\bar{k}) + \cos(k-\bar{k})](\delta_{k,\pi-\bar{k}} - \delta_{k,\bar{k}}) \\ &= -\frac{2g}{N}[1+\cos(2\bar{k})] \quad ,\end{aligned} \tag{3.72}$$

$$\begin{aligned}\langle \bar{k}, \pi-\bar{k}|\, S_B\, |\bar{k}, \pi-\bar{k}\rangle &= +\frac{g}{N}\,\langle \bar{k}, \pi-\bar{k}| \sum_{k} f_k^\dagger f_{\pi-k}^\dagger f_{\pi-k} f_k (1+\cos(2k))\, |\bar{k}, \pi-\bar{k}\rangle \\ &= -\frac{g}{N}\,\langle \bar{k}, \pi-\bar{k}| \sum_{k} f_k^\dagger f_{\pi-k}^\dagger (1+\cos(2k))(\delta_{k,\pi-\bar{k}} - \delta_{k,\bar{k}})\, |0\rangle \\ &= +\frac{g}{N}\,\langle \bar{k}, \pi-\bar{k}|\, \big[f_{\pi-\bar{k}}^\dagger f_{\bar{k}}^\dagger - f_{\bar{k}}^\dagger f_{\pi-\bar{k}}^\dagger\big](1+\cos(2\bar{k}))\, |0\rangle \\ &= +\frac{g}{N}(1+\cos(2\bar{k}))\,\langle \bar{k}, \pi-\bar{k}|\, \big[\, |\bar{k}, \pi-\bar{k}\rangle - |\pi-\bar{k}, \bar{k}\rangle\,\big] \\ &= +\frac{2g}{N}[1+\cos(2\bar{k})] \quad .\end{aligned} \tag{3.73}$$

Clearly, the two terms cancel each other and

$$\langle \bar{k}, \pi - \bar{k}| \, S_{N\%2} \, |\bar{k}, \pi - \bar{k}\rangle = 0 \quad . \tag{3.74}$$

3.4 Applicability of perturbation theory

Let us make use of this section by exploring in more detail the applicability of the perturbative approach, along with its possible problems. The first thing we will examine is how the method should be valid in case of small excitations. Then, we will check the regularity of the generating Hamiltonian, i.e. the one that creates the change of variables we used to find the first order normal form of the perturbation. After that, we will shortly discuss another setting in which the perturbative method just used — Jordan-Wigner change of coordinates, first order normal form upon unperturbed XY model — might work as well, even though we lose part of the symmetry of the system.

3.4.1 Small excitations

It has been briefly mentioned that our perturbative approach is justified in case of small excitations, or small M. Indeed, we are certainly not the first ones to consider $\hat{H}_z$ a perturbation: Davydov in his book [26] states that it can be treated as a perturbation whenever the number of excitations is small, i.e.

$$\frac{1}{N}\sum_n f_n^\dagger f_n = \frac{1}{N}\sum_k f_k^\dagger f_k << 1 \quad , \tag{3.75}$$

which corresponds to few spins being overturned. We can see from the N odd case that the validity of perturbation theory for small excitations is confirmed just by looking at the eigenvalues. We don't need to be concerned about the even case, as it only adds a term that does not really change the trend of the perturbation size with respect to M. Obviously, the unperturbed term gives an energy that scales proportionally to M:

$$E_0(k_1, ..., k_M) = g\Big[M - \sum_{j=1}^{M} \cos k_j\Big] \quad , \tag{3.76}$$

while the first order normal form of the perturbation actually goes like M^2/N, since it reads

$$E_1(k_1, ..., k_M) = -\frac{g}{N}\Big[M^2 - M - 2\sum_i \sum_{j,j>i} \cos(k_i - k_j)\Big] \quad . \tag{3.77}$$

Defining the ratio *perturbed-over-unperturbed*,

$$\rho(k_1, ..., k_M) := \left| \frac{E_1(k_1, ..., k_M)}{E_0(k_1, ..., k_M)} \right| \quad , \tag{3.78}$$

we find that, as M is an integer, the trend of ρ is roughly $\frac{M}{N}$. This clearly states that our perturbation is small enough for our purpose whenever it happens that $\frac{M}{N} << 1$. In particular, M could actually be a fraction of the number of sites, namely $M = \alpha N$, as long as $\alpha << 1$. Naturally, this becomes more and more possible for bigger values of N.

3.4.2 Generating Hamiltonian

In the procedure used to find the normal form of the perturbation, formally and completely described in appendix A, the change of variables that we implicitly performed is actually a canonical transformation which takes the form of a composition with the flow at a certain time of a certain Hamiltonian G, called the *generating Hamiltonian*. The further we go with the orders at which we want the perturbative Hamiltonian to be in normal form, the more of these generating Hamiltonians G_j we will have to use in the change of coordinates, which will look like a composition of flows (at times $\lambda, \lambda^2, ..., \lambda^j$ for each generator $G_1, G_2, ..., G_j$). These generating Hamiltonians are the actual unknowns of the perturbative construction, and have to be found order by order. By the nature of this construction, one has to make sure that the change of variables is actually achievable; such information is encoded into each generating Hamiltonian, under the regularity of the function. In other words, for us to be at ease with the procedure we require that the generating Hamiltonians do not diverge.

Let us see if, or when, this is true in our case. Since we truncated our Hamiltonian at the first order normal form, we only have to study the divergence of G_1, which, according to the theory, has the following form:

$$\hat{G_1} = \hat{\mathcal{G}}_1 + \frac{1}{T} \int_0^T (s - T)\, \delta \hat{P}_1 \circ \Phi_h^s \, \mathrm{d}s \quad , \tag{3.79}$$

where $\hat{\mathcal{G}}_1$ is an arbitrary element of $\ker(L_h)$, i.e. any operator such that $[\hat{\mathcal{G}}_1, \hat{h}] = 0$.

Recall that when we calculated $\langle \hat{P}_1 \rangle_h$, the result was a new condition, attached to the term in form of a Kronecker delta that involved the cosines of the momenta in the sum. Then, clearly, $\delta \hat{P}_1$ will be made of the exact same summation term, but in which the condition will be expressly the opposite

of what we had found. Therefore, defining

$$\Omega := \cos(k) + \cos(k') - \cos(q) - \cos(q') \quad , \tag{3.80}$$

we find that $\delta\hat{P}_1 = \hat{P}_1 - \langle\hat{P}_1\rangle_h$ reads

$$\begin{aligned}\delta\hat{P}_1 &= -\frac{g}{N}\sum_{\substack{k\,k'\,q\,q'\\ \Omega\neq 0}} f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')}\big(\delta_{k+q,k'+q'} + \delta_{k+q,k'+q'+2\pi} + \delta_{k+q,k'+q'-2\pi}\big) \\ &= -\frac{g}{N}\sum_{\substack{k\,k'\,q\,q'\\ \Omega\neq 0}} A(k,k',q,q')\Delta_{k+q,k'+q'} \quad ,\end{aligned} \tag{3.81}$$

where

$$A(k,k',q,q') = f_k^\dagger f_{k'} f_q^\dagger f_{q'} e^{-i(q-q')} \quad , \tag{3.82}$$

and

$$\Delta_{k+q,k'+q'} = \delta_{k+q,k'+q'} + \delta_{k+q,k'+q'+2\pi} + \delta_{k+q,k'+q'-2\pi} \quad . \tag{3.83}$$

Let us now build the whole computation for (3.79) piece by piece. First, let us define the primitive function

$$\begin{aligned}g(s) &:= \int \delta\hat{P}_1 \circ \Phi_h^s \,\mathrm{d}s = \int -\frac{g}{N}\sum_{\substack{k\,k'\,q\,q'\\ \Omega\neq 0}} A(k,k',q,q')\Delta_{k+q,k'+q'} e^{i\Omega s}\,\mathrm{d}s = \\ &= -\frac{g}{N}\sum_{\substack{k\,k'\,q\,q'\\ \Omega\neq 0}} A(k,k',q,q')\Delta_{k+q,k'+q'}\int e^{i\Omega s}\,\mathrm{d}s = \\ &= -\frac{g}{N}\sum_{\substack{k\,k'\,q\,q'\\ \Omega\neq 0}} A(k,k',q,q')\Delta_{k+q,k'+q'}\frac{e^{i\Omega s}}{i\Omega} \quad .\end{aligned} \tag{3.84}$$

We now use this to compute the following integral:

$$\begin{aligned}f(T) &:= \int_0^T (s-T)\,\delta\hat{P}_1 \circ \Phi_h^s \,\mathrm{d}s = \\ &= \left[(s-T)\int \delta\hat{P}_1 \circ \Phi_h^s \,\mathrm{d}s\right]_0^T - \int_0^T \delta\hat{P}_1 \circ \Phi_h^s \,\mathrm{d}s = \\ &= \Big[(s-T)g(s)\Big]_0^T = -Tg(T) =\end{aligned}$$

$$= +T\frac{g}{N}\sum_{\substack{k\,k'\,q\,q'\\ \Omega\neq 0}} A(k,k',q,q')\Delta_{k+q,k'+q'}\frac{1}{i\Omega} \quad . \tag{3.85}$$

Substituting this $f(T)$ into (3.79), we finally find

$$\hat{G}_1 = \hat{\mathcal{G}}_1 + \frac{g}{N}\sum_{\substack{k\,k'\,q\,q'\\ \Omega\neq 0}} A(k,k',q,q')\Delta_{k+q,k'+q'}\frac{1}{i\Omega} \quad . \tag{3.86}$$

Since both the deltas and the $A(k,k',q,q')$ are regular, the only possible source of divergence is the Ω at the denominator. This means that the generating Hamiltonian blows up whenever $\Omega = 0$, making our change of variables at least unreliable, if not even impossible. Fortunately we see, from the condition coming form $\delta\hat{P}_1$, that all combinations of momenta for which $\Omega = 0$ are not included in the sum. This leads us to think that we have nothing to worry about. It may very well be the case, as long as N is finite. Indeed, for finite values of N we have that any other combination of momenta yields an Ω which is definitely non-zero, as all the cosine values are well separated even for close sets of m_j such that $k_j = \frac{2\pi m_j}{N}$. Alas, this does not hold once we take the thermodynamic limit. The bigger N is, the closer together can be the values of momenta that we work with. Now, if we are only dealing with an issue like $k \sim k'$, $q \sim q'$, with each momenta far away from π or from 0, then we have that each difference of cosines goes like $\sim \frac{1}{N}$, which is taken care of by the factor $\frac{1}{N}$ in front of the sum, and there is no divergence. In fact, if we say $k = \epsilon \sim \frac{1}{N}$, i.e. $k = \frac{2\pi j}{N}$ with j such that $\lim_{N\to\infty}\frac{j}{N} = 0$, we explicitly have that

$$\cos(k) + \cos(q) - \cos(k+\epsilon) - \cos(q+\epsilon) = [\sin(k) + \sin(q)]\,\epsilon \sim \frac{1}{N} \quad . \tag{3.87}$$

However, if both couples are close together while also close to π or 0, the sum of sines in (3.87) is $\frac{1}{N}$ itself, so that the difference of cosines goes like $\frac{1}{N^2}$. This creates a divergence that cannot be removed. One can study it further and see that, typically, combinations of momenta which fluctuate closely about π and 0, or 2π naturally, do present a problem. This is just a quick assessment on the issues regarding this limit; the study of the thermodynamic limit usually requires very detailed examination, and we will exempt from it in this dissertation.

3.4.3 The parameter problem

As we said, in this chapter of the dissertation we have acted deliberately carefree of the prime concern of perturbation theory (of any kind); as it is stated in appendix A, a proper application of Hamiltonian perturbation theory requires the possibility to express the perturbation in implicit or explicit terms of a parameter that must be small ($\lambda << 1$). Our perturbation does definitely not have such an explicit dependence. In such cases, to prove that the theory is well suited for the system, one would have to find some norm $||\cdot||$ such that, if we have the Hamiltonian as

$$\hat{H} = \hat{h} + \hat{P}\ , \quad \hat{h} \text{ integrable}\ , \tag{3.88}$$

the ratio of the two Hamiltonian vector fields is small, i.e. $||X_P||/||X_h|| << 1$. This does not apply to us — our "perturbative" term $\hat{H}_z$ is most likely *not* a perturbation, in general, when part of the isotropic Heisenberg spin 1/2 chain—, but it works only in the case of a few overturned spins, or small excitations. Therefore, in addition to the possible problems at $N \to \infty$ expressed in 3.4.2, we might have more general problems of applicability if we consider generic states. In fact, the case of few overturned spins, while suitable for the perturbative treatment, is a restriction on the kind of *states* that can be studied with the perturbative approach, and not on the whole system — as opposed to whenever we restrict the applicability of perturbation theory for certain (small) values of the parameter λ, which defines the system itself, not the states. But perhaps something of this kind can be done here too, even though technically outside of the isotropic spin chain environment; it is suggested by the fact that the core of our perturbation $\hat{H}_z$ comes from changing variables on the z interaction term of the original Hamiltonian. Indeed, by inserting a parameter λ in front of the $\hat{S}^z_n S^z_{n+1}$ term, we are gifted with an explicit dependence of $\hat{H}_z$ on λ, along with its disappearance when $\lambda \to 0$. Obviously, this would mean breaking isotropy along the z axis, resulting in the XXZ chain.

3.4.4 The XXZ chain

Even though inserting the parameter in front of the z interaction implies inserting it in front of the term we called $\hat{H}_z$, this is not the end of the story. We recall that the change of variables of the z term also resulted in a quadratic term, which joined the other ones coming from the x and y terms into creating the unperturbed Hamiltonian $\hat{h}$. Let us quickly retrace the

steps we did, while we insert the new parameter λ:

$$\begin{aligned}\hat{H}_{XXZ} &= -g\sum_n \left[\frac{1}{2}\left(S_n^+ S_{n+1}^- + S_n^- S_{n+1}^+\right) + \lambda S_n^z S_{n+1}^z\right] = \\ &= -g\sum_n \left[\frac{1}{2}\left(f_n^\dagger f_{n+1} + f_n^\dagger f_{n-1}\right) - \lambda f_n^\dagger f_n + \lambda f_n^\dagger f_n f_{n+1}^\dagger f_{n+1}\right] \quad ,\end{aligned} \tag{3.89}$$

and under Fourier transform, we have

$$\hat{H}_{XXZ} = -g\sum_k \left[\cos k f_k^\dagger f_k - \lambda f_k^\dagger f_k + \lambda \hat{H}_z\right] \quad . \tag{3.90}$$

Since $\sum_k \lambda f_k^\dagger f_k$ has its time average equal to itself (recall that $\sum_k f_k^\dagger f_k$ is conserved), we have that there is nothing computationally different from the isotropic chain in evaluating the first order normal form. Therefore, the normal form for the XXZ chain at first order reads:

$$\hat{H}'_{XXZ} = \sum_k \tilde{\omega}(k,\lambda) f_k^\dagger f_k + \lambda \hat{S} \quad , \tag{3.91}$$

where

$$\tilde{\omega}(k,\lambda) = g(\lambda - \cos k) \quad . \tag{3.92}$$

3.5 N=6 example

Because of the difficulties explored in Subsection 3.3.3 in terms of generalizing the description of eigenvalue corrections for the N even case, it is interesting to witness what actually happens in those situations with an example. The simplest non-trivial case is for $N = 6$. The values of the momenta, which we will label with $1, 2, ..., 6$ are

j	1	2	3	4	5	6
k_j	$\frac{\pi}{3}$	$\frac{2\pi}{3}$	π	$\frac{4\pi}{3}$	$\frac{5\pi}{3}$	2π

,

while the important couplings for each kind of degeneracy are as follows:

$$\pi - k \longrightarrow 3 - j \qquad , \qquad 2\pi - k \longrightarrow 6 - j$$

j	$3-j$	$6-j$
1	2	5
2	1	4
3	6	3
4	5	2
5	4	1
6	3	6

.

3.5.1 Degeneracy for fixed M

The other kind of degeneracy, i.e. adding a 2π-momentum to any state, is here forgotten, as it actually has no coupling whatsoever — the two states have different M, which is a conserved quantity of the system. We will therefore discuss the degeneracy of the states for each M.

One excitation

Let us begin with the study of one excitation states, by listing the states and all the ones degenerate to them:

state	degenerate
$\|1\rangle$	$\|5\rangle$
$\|2\rangle$	$\|4\rangle$
$\|3\rangle$	/
$\|4\rangle$	$\|2\rangle$
$\|5\rangle$	$\|1\rangle$
$\|6\rangle$	/

(3.93)

So the degenerate subspaces are

$$\left\{ \begin{array}{c} |1\rangle \\ |5\rangle \end{array} \right\} , \left\{ \begin{array}{c} |2\rangle \\ |4\rangle \end{array} \right\} , \left\{ \; |3\rangle \; \right\} , \left\{ \; |6\rangle \; \right\} . \tag{3.94}$$

Everything is diagonal and there is no pair, so the only contributions are

$$\langle j| \, S_1 \, |j\rangle = 0 \quad , \tag{3.95}$$

so there is no correction to the unperturbed eigenvalues, which are 0, $\frac{1}{2}g$ (twice), $\frac{3}{2}g$ (twice), $2g$.

Two excitations

We do the same process for two excitations states $|j_1, j_2\rangle$, simplifying the degeneration table by not mentioning in the first column eigenstates that are already found degenerate to others. We denote with a star * the states which are paired by the nasty term of the Hamiltonian, i.e. the ones that will give birth to a non-diagonal matrix that will have to be diagonalized in order to find the eigenvalue corrections at first order.

state	degenerate
$\|1,2\rangle^*$	$\|2,5\rangle, \|1,4\rangle, \|3,6\rangle^*, \|4,5\rangle^*$
$\|1,3\rangle$	$\|3,5\rangle$
$\|1,5\rangle$	/
$\|1,6\rangle$	$\|5,6\rangle$
$\|2,3\rangle$	$\|3,4\rangle$
$\|2,4\rangle$	/
$\|2,6\rangle$	$\|4,6\rangle$

(3.96)

The degenerate subspaces are:

$$\left\{ \begin{array}{l} |1,2\rangle^* \\ |1,4\rangle \\ |2,5\rangle \\ |3,6\rangle^* \\ |4,5\rangle^* \end{array} \right\} , \left\{ \begin{array}{l} |1,3\rangle \\ |3,5\rangle \end{array} \right\} , \left\{ \begin{array}{l} |1,6\rangle \\ |5,6\rangle \end{array} \right\} , \left\{ \begin{array}{l} |2,3\rangle \\ |3,4\rangle \end{array} \right\} , \left\{ \begin{array}{l} |2,6\rangle \\ |4,6\rangle \end{array} \right\} ,$$

$$\left\{ \ |1,5\rangle \ \right\} , \left\{ \ |2,4\rangle \ \right\} . \tag{3.97}$$

Within each space we study the interaction matrix. All the diagonal terms are, as usual,

$$\langle j_1, j_2| \, S_1 \, |j_1, j_2\rangle = -\frac{2g}{N}[1 - \cos(k_{j_1} - k_{j_2})] \quad , \tag{3.98}$$

while the non-diagonal terms are zero whenever one of the states is not denoted by a star *. Whenever they are both starred, we get

$$\langle j_1, 3 - j_1| \, S_{N\%2} \, |j_2, 3 - j_2\rangle = -\frac{g}{N} \cos(k_{j_1} + k_{j_2}) \quad . \tag{3.99}$$

We could compute here the corrections for each subspace, although it is clear that in those spaces where we have diagonal interaction we will find that all the corrections are the same, and there is nothing interesting about it. Thus,

we will skip those calculations and focus on the subspace

$$\left\{ \begin{array}{c} |1,2\rangle^* \\ |1,4\rangle \\ |2,5\rangle \\ |3,6\rangle^* \\ |4,5\rangle^* \end{array} \right\} , \quad \text{with eigenvalue } \; E^{(0)} = 2g \; .$$

Using formulas (3.98) and (3.99), we find the interaction matrix to be

$$\mathcal{S} = \begin{bmatrix} -4\frac{g}{N} & 0 & 0 & 0 & 0 \\ 0 & -4\frac{g}{N} & 0 & 0 & 0 \\ 0 & 0 & -\frac{g}{N} & \frac{g}{2N} & -\frac{g}{2N} \\ 0 & 0 & \frac{g}{2N} & -4\frac{g}{N} & -\frac{g}{2N} \\ 0 & 0 & -\frac{g}{2N} & -\frac{g}{2N} & -\frac{g}{N} \end{bmatrix} \quad . \tag{3.100}$$

As we expected, we are presented with a diagonal block, corresponding to those state with no pair, and a non diagonal one, which is diagonalized in the following:

$$-\frac{g}{12}\begin{bmatrix} 2 & -1 & 1 \\ -1 & 8 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \longrightarrow \quad -\frac{g}{12}\begin{bmatrix} 3 & 0 & 0 \\ 0 & \frac{9+\sqrt{57}}{2} & 0 \\ 0 & 0 & \frac{9-\sqrt{57}}{2} \end{bmatrix} \quad . \tag{3.101}$$

Thus, the five new eigenvalues read

$$E^{(0)} + E^{(1)} = \begin{cases} \frac{4}{3}g \text{ (twice)} \\ \frac{7}{4}g \\ \frac{39+\sqrt{57}}{24}g \\ \frac{39-\sqrt{57}}{24}g \end{cases} \quad . \tag{3.102}$$

Three excitations

The case of three excitations in the most cumbersome in the $N = 6$ case, with four subspaces that have non-diagonal interaction matrices. Let us do the same as we did in the previous subsection and classify the states that are degenerate, this time with two different notations for pairing, one star * and two **, as there are cases where in a single subspace there are more than one sub-subspace that pair. This will give rise to block matrices. The grouping

is the following:

state	degenerate
$\|1,2,3\rangle^*$	$\|2,3,5\rangle$, $\|1,3,4\rangle$, $\|3,4,5\rangle^*$
$\|1,2,4\rangle^*$	$\|2,4,5\rangle^{**}$, $\|3,4,6\rangle^*$, $\|2,3,6\rangle^{**}$
$\|1,2,5\rangle^*$	$\|1,4,5\rangle^{**}$, $\|3,5,6\rangle^*$, $\|1,3,6\rangle^{**}$
$\|1,2,6\rangle^*$	$\|2,5,6\rangle$, $\|1,4,6\rangle$, $\|4,5,6\rangle^*$
$\|1,3,5\rangle$	/
$\|1,5,6\rangle$	/
$\|2,3,4\rangle$	/
$\|2,4,6\rangle$	/

(3.103)

so the degenerate subspaces are:

$$\left\{ \begin{array}{l} |1,2,3\rangle^* \\ |2,3,5\rangle \\ |1,3,4\rangle \\ |3,4,5\rangle^* \end{array} \right\} , \left\{ \begin{array}{l} |1,2,4\rangle^* \\ |2,4,5\rangle^{**} \\ |3,4,6\rangle^* \\ |2,3,6\rangle^{**} \end{array} \right\} , \left\{ \begin{array}{l} |1,2,5\rangle^* \\ |1,4,5\rangle^{**} \\ |3,5,6\rangle^* \\ |1,3,6\rangle^{**} \end{array} \right\} , \left\{ \begin{array}{l} |1,2,6\rangle^* \\ |2,5,6\rangle \\ |1,4,6\rangle \\ |4,5,6\rangle^* \end{array} \right\} ,$$

$$\left\{ \ |1,3,5\rangle \ \right\} , \left\{ \ |1,5,6\rangle \ \right\} , \left\{ \ |2,3,4\rangle \ \right\} , \left\{ \ |2,4,6\rangle \ \right\} . \qquad (3.104)$$

Let us go through each type of term possible in this case. All the diagonal terms are, as usual,

$$\langle j_1, j_2, j_3| \, S_1 \, |j_1, j_2, j_3\rangle = -\frac{2g}{N}[3 - \cos(k_{j_1} - k_{j_2}) - \cos(k_{j_1} - k_{j_3}) - \cos(k_{j_2} - k_{j_3})] \quad , \qquad (3.105)$$

while the non-diagonal terms are zero whenever one of the states is not denoted by some kind of star *. Whenever they are both starred, we get

$$\langle j_1, 3 - j_1, l_1| \, S_{N\%2} \, |j_2, 3 - j_2, l_2\rangle = -\frac{g}{N} \cos(k_{j_1} + k_{j_2}) \delta_{l_1,l_2} \quad . \qquad (3.106)$$

That delta, then, indicates that not all the couples of states with a pair inside have nonzero interaction. This is precisely the reason why we starred differently some of the states: to indicate which ones actually interact with each other. It turns out that the spaces that have the same coupling (i.e. either two single states and two coupled or four states coupled two by two) give exactly the same first order corrections. For this reason we will show the diagonalization of only two of them.

Let us start with the subspace

$$\left\{ \begin{array}{l} |1,2,3\rangle^* \\ |2,3,5\rangle \\ |1,3,4\rangle \\ |3,4,5\rangle^* \end{array} \right\} , \quad \text{with eigenvalue } E^{(0)} = 4g \ . \qquad (3.107)$$

Using formulas (3.105) and (3.106), we find the interaction matrix to be

$$\mathcal{S} = \begin{bmatrix} -8\frac{g}{N} & 0 & 0 & 0 \\ 0 & -8\frac{g}{N} & 0 & 0 \\ 0 & 0 & -5\frac{g}{N} & -\frac{g}{2N} \\ 0 & 0 & -\frac{g}{2N} & -5\frac{g}{N} \end{bmatrix} \quad . \tag{3.108}$$

Just like with two excitations, we are presented with a diagonal block and a non diagonal one, the latter being is diagonalized in the following:

$$-\frac{g}{12}\begin{bmatrix} 10 & 1 \\ 1 & 10 \end{bmatrix} \quad \longrightarrow \quad -\frac{g}{12}\begin{bmatrix} 11 & 0 \\ 0 & 9 \end{bmatrix} \quad . \tag{3.109}$$

Thus, the four new eigenvalues read

$$E^{(0)} + E^{(1)} = \begin{cases} \frac{5}{3}g \text{ (twice)} \\ \frac{25}{12}g \\ \frac{9}{4}g \end{cases} \quad . \tag{3.110}$$

We now work with the other kind of subspaces, for example

$$\left\{ \begin{array}{l} |1,2,4\rangle^{*} \\ |2,4,5\rangle^{**} \\ |3,4,6\rangle^{*} \\ |2,3,6\rangle^{**} \end{array} \right\} , \quad \text{with eigenvalue } E^{(0)} = \frac{7}{2}g \, . \tag{3.111}$$

Using formulas (3.105) and (3.106), we find the interaction matrix to be

$$\mathcal{S} = \begin{bmatrix} -8\frac{g}{N} & -\frac{g}{2N} & 0 & 0 \\ -\frac{g}{2N} & -8\frac{g}{N} & 0 & 0 \\ 0 & 0 & -8\frac{g}{N} & \frac{g}{2N} \\ 0 & 0 & -\frac{g}{2N} & -8\frac{g}{N} \end{bmatrix} \quad . \tag{3.112}$$

This time, we have two non diagonal blocks, both of which though have the same eigenvalues

$$-\frac{g}{12}\begin{bmatrix} 16 & \pm 1 \\ \pm 1 & 16 \end{bmatrix} \quad \longrightarrow \quad -\frac{g}{12}\begin{bmatrix} 17 & 0 \\ 0 & 15 \end{bmatrix} \quad . \tag{3.113}$$

Thus, the four new eigenvalues read

$$E^{(0)} + E^{(1)} = \begin{cases} \frac{25}{12}g \text{ (twice)} \\ \frac{9}{4}g \text{ (twice)} \end{cases} \quad . \tag{3.114}$$

Four excitations

We do the same for $M = 4$. Now, in this case, any state will contain at least one pair. This doesn't necessarily mean that it interacts with other state; in fact, we will see that the only interacting states are the ones with two pairs inside, which will be signaled with the usual star *. Once again we have the table of degeneracies:

state	degenerate
$\|1,2,3,4\rangle$	$\|2,3,4,5\rangle$
$\|1,2,3,5\rangle$	$\|1,3,4,5\rangle$
$\|1,2,3,6\rangle^*$	$\|2,3,5,6\rangle$, $\|1,3,4,6\rangle$, $\|3,4,5,6\rangle^*$, $\|1,2,4,5\rangle^*$
$\|1,2,4,6\rangle$	$\|2,4,5,6\rangle$
$\|1,2,5,6\rangle$	$\|1,4,5,6\rangle$
$\|1,3,5,6\rangle$	/
$\|2,3,4,6\rangle$	/

(3.115)

The degenerate subspaces are:

$$\left\{ \begin{array}{l} |1,2,3,6\rangle^* \\ |2,3,5,6\rangle \\ |1,3,4,6\rangle \\ |3,4,5,6\rangle^* \\ |1,2,4,5\rangle^* \end{array} \right\} , \left\{ \begin{array}{l} |1,2,3,4\rangle \\ |2,3,4,5\rangle \end{array} \right\} , \left\{ \begin{array}{l} |1,2,3,5\rangle \\ |1,3,4,5\rangle \end{array} \right\} , \left\{ \begin{array}{l} |1,2,4,6\rangle \\ |2,4,5,6\rangle \end{array} \right\} ,$$
$$\left\{ \begin{array}{l} |1,2,5,6\rangle \\ |1,4,5,6\rangle \end{array} \right\} , \{ \, |1,3,5,6\rangle \, \} , \{ \, |2,3,4,6\rangle \, \} \, . \tag{3.116}$$

Once again, the diagonal terms will read:

$$\begin{aligned} \langle j_1, j_2, j_3, j_4 | \, S_1 \, | j_1, j_2, j_3, j_4 \rangle = -\frac{2g}{N} [6 - \cos(k_{j_1} - k_{j_2}) - \cos(k_{j_1} - k_{j_3}) - \\ - \cos(k_{j_2} - k_{j_3}) - \cos(k_{j_1} - k_{j_4}) - \\ - \cos(k_{j_2} - k_{j_4}) - \cos(k_{j_3} - k_{j_4})] \quad , \end{aligned} \tag{3.117}$$

while the non-diagonal terms are zero whenever one of the states is not denoted by a star * (you can easily check that all others are zero). Whenever they are both starred, we get — notice that in couples of such states, there is always one pair $\{k, \pi - k\}$ which is the same in both —

$$\langle j_1, 3 - j_1, j, 3 - j | \, S_{N\%2} \, | j_2, 3 - j_2, j, 3 - j \rangle = -\frac{g}{N} \cos(k_{j_1} + k_{j_2}) \quad . \tag{3.118}$$

Let us then diagonalize the subspace

$$\left\{ \begin{array}{l} |1,2,3,6\rangle^* \\ |2,3,5,6\rangle \\ |1,3,4,6\rangle \\ |3,4,5,6\rangle^* \\ |1,2,4,5\rangle^* \end{array} \right\} , \quad \text{with eigenvalue } E^{(0)} = 4g \, .$$

Using formulas (3.117) and (3.118), we find the interaction matrix to be

$$\mathcal{S} = \begin{bmatrix} -16\frac{g}{N} & 0 & 0 & 0 & 0 \\ 0 & -16\frac{g}{N} & 0 & 0 & 0 \\ 0 & 0 & -13\frac{g}{N} & -\frac{g}{2N} & -\frac{g}{2N} \\ 0 & 0 & -\frac{g}{2N} & -13\frac{g}{N} & \frac{g}{2N} \\ 0 & 0 & -\frac{g}{2N} & \frac{g}{2N} & -16\frac{g}{N} \end{bmatrix} \quad . \tag{3.119}$$

Once again, we have a diagonal block, corresponding to those state that have no interaction, and a non diagonal one, which is diagonalized in the following:

$$-\frac{g}{12}\begin{bmatrix} 26 & 1 & 1 \\ 1 & 26 & -1 \\ 1 & -1 & 32 \end{bmatrix} \quad \longrightarrow \quad -\frac{g}{12}\begin{bmatrix} 27 & 0 & 0 \\ 0 & \frac{57+\sqrt{57}}{2} & 0 \\ 0 & 0 & \frac{57-\sqrt{57}}{2} \end{bmatrix} \quad . \tag{3.120}$$

Thus, the five new eigenvalues read

$$E^{(0)} + E^{(1)} = \begin{cases} \frac{4}{3}g \text{ (twice)} \\ \frac{7}{4}g \\ \frac{39+\sqrt{57}}{24}g \\ \frac{39-\sqrt{57}}{24}g \end{cases} \quad . \tag{3.121}$$

Remarkably, we found the same final eigenvalues as the $M = 2$ case. This is not only true for the conjugated states; as one can easily prove, the complete set of $M = 2$ eigenstates matches with the $M = 4$ one.

Five excitations

It gets easier now with $M = 5$, and with same properties of degeneracy as the $M = 1$ case:

state	degenerate
$\|1,2,3,4,5\rangle$	/
$\|1,2,3,4,6\rangle$	$\|2,3,4,5,6\rangle$
$\|1,2,3,5,6\rangle$	$\|1,3,4,5,6\rangle$
$\|1,2,4,5,6\rangle$	/

, (3.122)

with the degenerate subspaces (all with diagonal interaction)

$$\left\{ \begin{array}{l} |1,2,3,4,6\rangle \\ |2,3,4,5,6\rangle \end{array} \right\} , \left\{ \begin{array}{l} |1,2,3,5,6\rangle \\ |1,3,4,5,6\rangle \end{array} \right\} , \left\{ \; |1,2,3,4,5\rangle \; \right\} , \left\{ \; |1,2,4,5,6\rangle \; \right\} . \tag{3.123}$$

The diagonal interaction, thus the corrections, will be calculated this way:

$$\begin{aligned} \langle j_1,j_2,j_3,j_4,j_5|\, S_1 \,|j_1,j_2,j_3,j_4,j_5\rangle = -\frac{2g}{N}[10 &- \cos(k_{j_1}-k_{j_2}) - \cos(k_{j_1}-k_{j_3}) - \\ &- \cos(k_{j_2}-k_{j_3}) - \cos(k_{j_1}-k_{j_4}) - \\ &- \cos(k_{j_2}-k_{j_4}) - \cos(k_{j_3}-k_{j_4}) - \\ &- \cos(k_{j_1}-k_{j_5}) - \cos(k_{j_2}-k_{j_5}) - \\ &- \cos(k_{j_3}-k_{j_5}) - \cos(k_{j_4}-k_{j_5})] \quad , \end{aligned} \tag{3.124}$$

but there is nothing more to say about it, and the process is quite trivial. The corrections actually all give the same result ($-4g$), and the order of the energy states in herein preserved from the unperturbed one. Once again, though, we find that the final eigenvalues are the same as the $M = 1$ case, meaning 0, $\frac{1}{2}g$ (twice), $\frac{3}{2}g$ (twice), $2g$.

Six excitations

Even more trivial is, naturally, the $M = 6$ case, for which we have only one state

$$|1,2,3,4,5,6\rangle \;, \quad \text{with eigenvalue } E^{(0)} = 6g \;. \tag{3.125}$$

The correction, calculated the usual way, actually gives an interesting result: $-6g$. This means, that at first order approximation, our model in this state has the energy

$$E^{(0)} + E^{(1)} = 0 \quad . \tag{3.126}$$

As we found in the previous cases, this state goes, from the unperturbed system to the perturbed one, from one of the highest energy states to one with energy equal to the ground state.

3.5.2 Conclusions

What we found in this simple example, i.e. that the unperturbed higher energy state go back to lower levels once the perturbation is added, is actually

intuitively true. The $M = 6$ state, for example, has all the spins flipped, i.e. is completely polarized. It is natural that, in the case of no external magnetic field such as it is ours, this has the same energy as the ground state, which corresponds to all spins in the same direction, even though it is the opposite one — we do not really have a privileged direction here. The same goes for the pairs $M = 1, M = 5$ and $M = 2, M = 4$, as one is the "flipped" version of the other, i.e. have same, but opposite, total z-component. If there is no preferred z direction, like in our model, these pairs describe the same states, and therefore have the same energy. In fact, the difference in ordering the energies with M in the unperturbed case ($M = 0$ ground state, $M = 6$ highest possible state) is actually due to the fact that the unperturbed system is equivalent to the XY model, but with a constant external field which breaks the symmetry under parity (along z), making a completely polarized state in one direction much less likely than the one in the opposite direction.

Conceptually, this explanation is quite neat and understandable, although at a first glance it might seem mysterious how intuitively true our perturbed result looks to be. In fact, we are working in the isotropic case, where our approach is theoretically reliable only for really small values of the ratio $\frac{M}{N}$; instead, we have calculated the eigenvalues regardless of this requirement not being met, yet still found (apparently) sensible results — meaning, results that follow our intuitive understanding of the system. In reality, the ordering of M is restored because we have resumed the z inversion symmetry just by adding the first order normal form of the perturbation, while the actual values of the eigenstates might be very far from the true ones. Indeed, recalling that the one quadratic term coming from z-z interaction had its average equal to itself (see 3.4.4), we see that the only thing we did to the z-z interaction term $\hat{S}^z_n \hat{S}^z_{n+1}$ was performing the average upon the flow of the unperturbed system — an action which ultimately preserves the symmetry by spin flip.

3.6 Comparison with the Bethe Ansatz

Having estimated the spectrum resulting from the Hamiltonian in normal form (at least in a simple way for the N odd case), one would straight away consider studying a parallel with the energy levels found by the Bethe ansatz, the integrable technique that was first able to exactly solve the Heisenberg chain. Unfortunately, the comparison proved to be more of a challenge than what our expectations were, and we were not able to study it in the majority of the cases. On our part, the case of N even is difficult to approach because of the computation being particularly case specific, as shown in the previous section for the $N = 6$ case. For the N odd case, the difficulties have more

to do with the ansatz and its relationship with our method. The three main challenges are the following:

- **The complexity of the integrable technique and its solutions** — as one can see from appendix B, finding the complex solutions for the Bethe ansatz is generally not an analytical process, and mainly requires the use of numerical methods. Furthermore, the identification and classification of complex solutions still an open problem at $M > 2$ already (for a detailed account on such computational difficulties see [12]). What is usually done to solve the system is to resort to the so called *string hypothesis* — see appendix B. First of all, we are unsure if this approach actually does exhaust all the solutions; not to mention, the whole string hypothesis requires the thermodynamic limit — which, as we explained in Section 3.4.2, may present a problem in our approach.

- **Different variables** — while the notation might be misleading, the ks present in the Bethe approach are not the same as ours. This represents a problem when trying to check if the two formulas for the energies (ours and Bethe's) coincide, or in trying to assess in which regime they become similar, regardless of the computational difficulties in solving the Bethe equations. First, as one immediately sees from the $M = 2$ discussion in the appendix, the Bethe are the following

$$k_{1,2} = \frac{2\pi m_{1,2}}{N} \pm \frac{\theta}{N} \quad , \tag{3.127}$$

  where θ is a phase with the purpose of carrying information on the interaction between the two magnons. But even if the first part of the Bethe ks looks like our way of writing them, this is actually not true. In fact, the integers m there described do not have fermionic properties, and one can find solutions in which the two ks contain the same value of m — this is untrue for our variables. The reason for this can be found by looking at the way the wavefunction is built for the case $M = 1$ of the Bethe ansatz:

$$|\psi\rangle = \sum_n e^{ikn} |n\rangle = \sum_n e^{ikn} \hat{S}_n^+ |0\rangle \quad , \tag{3.128}$$

  which seems to correspond to a Fourier transform upon the ladder operators themselves, instead of upon the fermionic creation and annihilation operators we deal with. This recalls the *Holstein-Primakov transformation* — see [27] — for spin $1/2$, in which the ladder operators

directly identify creation and annihilation operators of a kind of particle that is not a fermion nor a boson (anti-commutation relations on same site, bosonic commutation relations on different sites). This aspect of the Bethe ansatz seems to indicate the Holstein-Primakov transformation as the most suitable change of variables to use if searching for a match with the energy levels found by the ansatz. The reason we did not use such a change of coordinates is precisely for the fact that it does not map the system into the analogue system made of known types of particles, bosons or fermions, but it maps it into a system of weird hybrid ones. One could try anyways and retrace the steps we did, substituting the Jordan-Wigner transformation with Holstein-Primakov, and see if a comparison with the Bethe ansatz is easier that way.

The problem of the non-fermionic properties of the integers m in the Bethe ansatz is actually removed in that instance by shifting to other variables — the *rapidities* λ —, but we remain unsure as to if a comparison is possible between our energy levels and the Bethe ones in terms of the rapidities; even though the rapidities seem to be fermions, it is quite difficult to understand, through the change of variables performed form k to λ, if these fermions are actually the same variables we end up with via the Jordan-Wigner transformation. In other words, our process and the Bethe process are different both in terms of approach and of variables involved, creating non-negligible complications when attempting to build a bridge between the two results.

A possible avenue, to be further explored at another time, would be to actually adapt the Bethe ansatz approach to the Heisenberg system already fermionized via Jordan-Wigner, and see if the Fourier-like coefficients of the wave functions can be more easily used for the comparison with the ks of the Fourier transform we used on our fermionic operators. More importantly, one would be interested in checking if the new variables provide some advantage to the computational complexity of the technique.

- **Approximate results vs. accurate results** — In addition to the two substantial issues above, we also have that while the energy levels found by Bethe are exact (whenever the solution is available), ours are certainly not, so there is never going to be an exact correspondence between the two. This is especially true in the case of a ratio M/N which is not small, since, as we know from Section 3.4, the system is not in a perturbative regime anymore, and our method fails.

3.6.1 Computational comparison: M=2

Despite the challenges presented, it is possible to get an idea of how the Bethe energie levels are arranged, and compare the trend with our results, at least in a simple case, that of $M = 2$. In that case, solving the Bethe equations is computationally approachable, especially once we notice that the Bethe equations in terms of rapidities for two variables give converging solutions using a simple algorithm for solving a system of nonlinear equation. The method is not 100% accurate, and might present some issues in a few cases; however, the general trend of the energy levels should be similar enough to the real solutions, with an average error that becomes smaller for bigger values of N. Comparing different cases of N, we confirm that our approach gets better at approximating the spectrum for smaller values of $\frac{M}{N}$, as we can see from the following graphs.

The case of $N = 6$ shows that while the trend looks similar, our results are still too approximated:

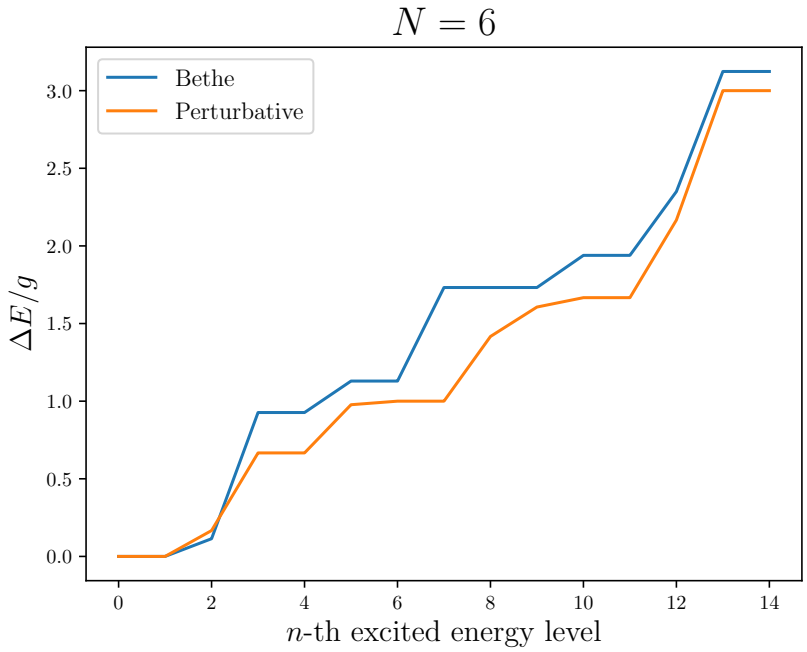


Figure 3.1: Comparison with Bethe ansatz for N=6

This is because in the case of $N = 6$ the parameter is $\frac{M}{N} = \frac{1}{3}$, which is not small enough, and not just because the algorithm is less reliable. In fact, in [12] we find a table for accurate energy levels of the Bethe ansatz for $N = 5$, $M = 2$; plotting them together with our results, we see that while the trend

is ok, the approziamtion is not very satisfactory:

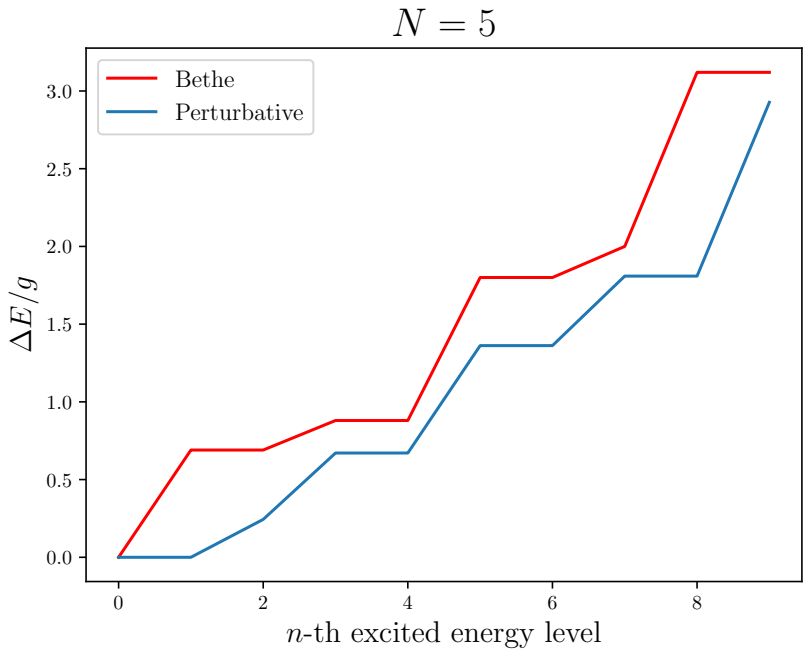


Figure 3.2: Comparison with Bethe ansatz (accurate) for N=5

The situation improves for bigger values of N, and we report here the graphs for $N = 25$ and $N = 101$, with parameters $\frac{M}{N}$ respectively 0.08 and 0.02 :

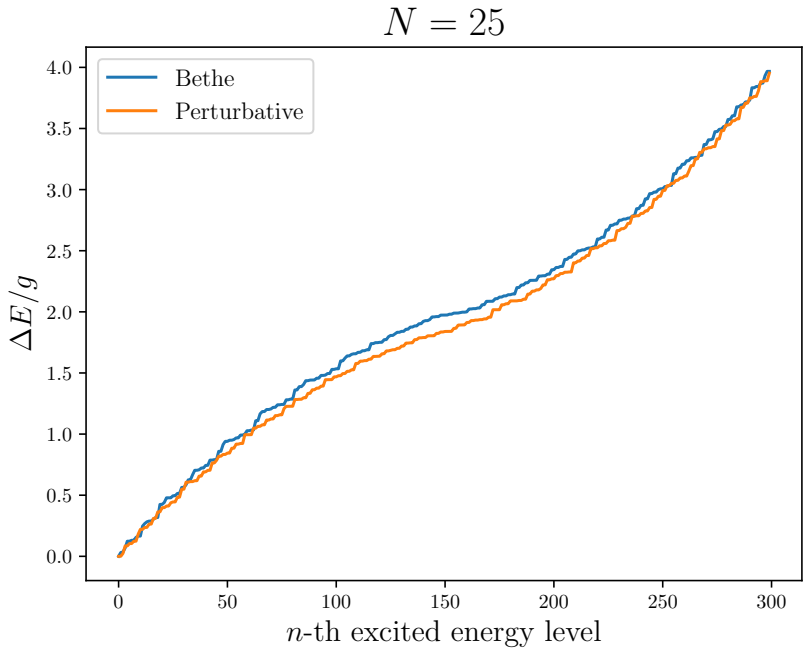


Figure 3.3: Comparison with Bethe ansatz for N=25

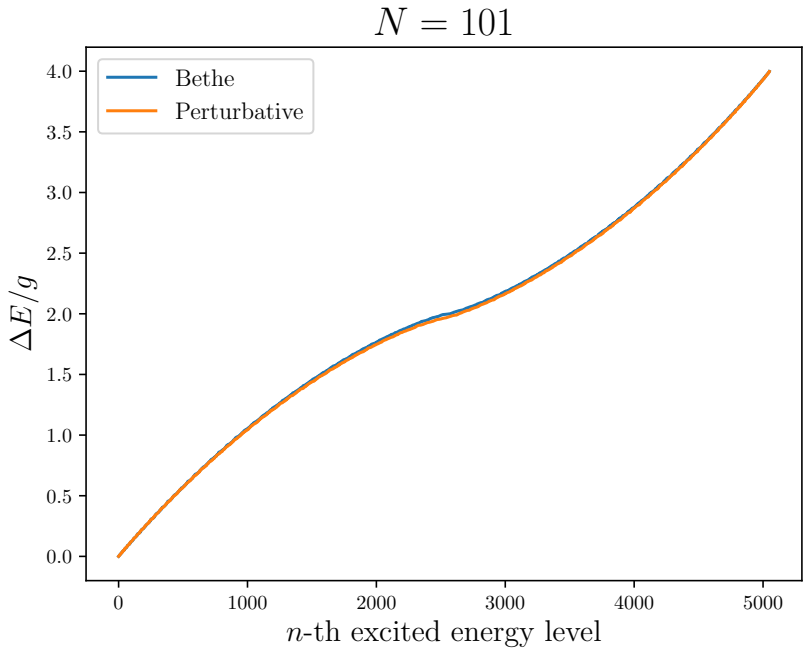


Figure 3.4: Comparison with Bethe ansatz for N=101

Clearly, as we can see, the smaller the parameter is the more correctly our method describes the system, just as we expected.

Appendix A

Hamiltonian perturbation theory

Perturbation theory was first introduced in the field of celestial mechanics, owing its name to the works of Lagrange and Laplace. It was later studied by great scholars — Poincarè, Birkhoff and Kolmogorov, the latter responsible for the well known KAM theorem for quasi-periodic motion. The general theory of non-linear oscillations was developed by Krylov and Bogolyubov. As one can see from the list of people involved, but also on the amount of works still on the topic, the approach has met great encouragement from the scientific community, to speak for its unprecedented success in the field [28].

In general, perturbation theory is based on the fact that it is possible to give an approximate description of a system by comparing it to an ideal and solvable one that is somehow close to it. One of the criteria of applicability of perturbation theory is that whatever differentiates the real system from the ideal one must be, implicitly or explicitly, dependent on some small parameter — and that whenever such parameter is zero, the system must revert to the ideal model. In quantum mechanics, perturbation theory takes the form of an eigenvalue correction for stationary states, thus retaining a static outlook at the system. Even so, it is widely used and is taught in every course on condensed matter theory.

What is proposed here is a slight generalization to classical Hamiltonian perturbation theory, exploiting the advantages of the Poisson algebra formulation of classical systems, [22], [21], which allows freedom of coordinate change. Actually, a Poisson algebra formulation is possible even in the case of quantum mechanics — Heisenberg's operator algebra is a suitable one — and we will see that the results of classical Hamiltonian perturbation theory can be translated in a quantum setting very easily. This result was also obtained in [29], restricted to Birkhoff-Gustavson normal forms, where a Lie algebra framework for the classical systems allows a parallel between classical and quantum normal forms.

A.1 Classical perturbation theory

We start by considering a Hamiltonian of the form

$$H_\lambda = h + \lambda P_1 + \lambda^2 P_2 + ... + \lambda^n P_n + R_{n+1} \ , \tag{A.1}$$

where λ is a small parameter ($|\lambda| << 1$), the functions P_j for $j = 1, ..., n$ are λ-independent, and $R_{n+1} = O(\lambda^{n+1})$.
If the system described by the Hamiltonian h is *integrable*, meaning that its flow $\Phi_h^t \quad \forall t \in \mathbb{R}$ is known, H_λ is said to be *quasi-integrable* or *close to integrable*, while $P(\lambda) = H_\lambda - h$ is called the *perturbation*, or *perturbative Hamiltonian*. Generally, the perturbation P does not appear explicitly ordered in terms of the parameter λ. In those cases, the Hamiltonian is then split into a leading part (P_1) plus a remainder (R_2) , such remainder in turn is then possibly split into a leading part (P_2) and so on. This splitting can be different in different regions of the phase space. When we split the perturbation like so, we can artificially insert the parameter λ as a *tracer* of the ordering, and we will set it to one at the end of the calculations. When we deal with perturbations that do not have any explicit small parameter ordering, we must be careful about the fact that the Hamiltonian P might actually not be fit to treat as a perturbation, meaning that it is not "small" compared to the unperturbed system. We must then study the closeness of the perturbed system H_λ to the unperturbed one defined by h. What really matters in determining it is not really the ratio $|P|/|h|$ — to see this, remember that a constant perturbation does not affect the dynamics independently of its size. More importantly, it is the ratio of the respective vector fields, i.e. $\|X_P\|/\|X_h\| << 1$, in some norm.

The procedure at the heart of Hamiltonian perturbation theory consists in looking for a change of variables which removes, completely or at least partially, the perturbation $P(\lambda)$ from the original Hamiltonian H_λ , up to a pre-fixed order. Unfortunately, in general, the complete removal of the perturbation, even just at first order, is not possible. Therefore, we resort to a partial removal of the perturbation (still, up to some order), i.e. a transformation of the perturbative Hamiltonian into another form — called *normal form* — which possesses useful properties. It is defined in the following:

Definition A.1.1. (**Normal form**) A Hamiltonian $\mathcal{H}_\lambda$ of the form

$$\mathcal{H}_\lambda = h + \sum_{j=1}^{n} \lambda^j S_j + R_{n+1} \ , \tag{A.2}$$

is said to be in *normal form up to order n with respect to h*, if $\{S_j, h\} = 0$ for any $j = 1, ..., n$ and $R_{n+1} = O(\lambda^{n+1})$.

We notice that the Hamiltonian (A.2) is of the form (A.1), but now the perturbation consists of the summation of first integrals S_j of h. Notice that this includes the case $S_j = 0$ for some j , i.e. the complete removal of the perturbation at the order j. So, from a technical point of view, Hamiltonian perturbation theory aims to find a suitable change of variables that maps the quasi-integrable Hamiltonian (A.1) into its normal form of the type (A.2) , i.e. $H_\lambda \to \mathcal{H}_\lambda$. Let us state this more precisely.

Let Γ be the phase space, D the definition domain of the Hamiltonian and let us suppose that

1. the Hamiltonian H_λ has the form (A.1) for any $x \in D \subset \Gamma$;
2. the unperturbed Hamiltonian h is integrable in D, i.e. its flow $\Phi_h^t(\xi)$, which is the solution of the Hamilton equations
$$\dot{x} = X_h(x) = J\nabla H(x),$$
is known for any $\xi \in D$;
3. the flow $\Phi_h^t(\xi)$ is bounded in D uniformly in time, i.e. there exists a constant C such that
$$\|\Phi_h^t(\xi)\| \leq C$$
for any $\xi \in D$ in some suitable norm $\|\cdot\|$.

With this, we look for a canonical transformation

$$\mathcal{C}_\lambda : x \mapsto y = \mathcal{C}_\lambda(x) = x + \mathcal{O}(\lambda) \ , \tag{A.3}$$

which we require to be *smooth*, λ-*dependent* and λ-*close to the identity*, such that

$$\mathcal{H}_\lambda(y) := H_\lambda(\mathcal{C}_\lambda^{-1}(y)) \tag{A.4}$$

is in normal form up to a certain order n with respect to h. The λ-closeness to the identity of this map is necessary in order to match the unperturbed problem as $\lambda \mapsto 0$. The canonicity of $\mathcal{C}_\lambda$ is instead optional; we only require it in order to perform the transformation of the Hamiltonian function without ever minding about the consequent deformation of the Poisson structure. However, this is just a concern for the sake of simplicity, a choice; unfortunately, it does produce a substantial restriction in our perturbation theory framework.

A.1.1 Construction of $\mathcal{C}_\lambda$ and the Averaging Principle

Our approach into building the above mentioned canonical transformation $\mathcal{C}_\lambda$, is to assume it to be a composition of Hamiltonian flows of suitable generating Hamiltonians, each at λ-dependent times.
We then look for a canonical transformation $\mathcal{C}_\lambda$ of the form

$$\mathcal{C}_\lambda = \mathcal{C}_n \circ \mathcal{C}_{n-1} \circ ... \circ \mathcal{C}_2 \circ \mathcal{C}_1 \tag{A.5}$$

with its inverse

$$\mathcal{C}_\lambda^{-1} = \mathcal{C}_1^{-1} \circ \mathcal{C}_2^{-1} \circ ... \circ \mathcal{C}_{n-1}^{-1} \circ \mathcal{C}_n^{-1} \quad , \tag{A.6}$$

where we choose

$$\mathcal{C}_j^{-1} = \Phi_{G_j}^{\lambda^j} \ , \quad \mathcal{C}_j = \Phi_{G_j}^{-\lambda^j} \ , \quad \forall j = 1, ..., n \quad , \tag{A.7}$$

and where the choice of the minus sign in front of λ^j in the direct transformation is just a convention. Here, $\Phi_{G_j}^{\pm\lambda^j}$ is the flow of a certain Hamiltonian G_j respectively at time $\pm\lambda^j$. Let us state the whole transformation explicitly:

$$\mathcal{C}_\lambda^{-1} = \Phi_{G_1}^{\lambda} \circ \Phi_{G_2}^{\lambda^2} \circ ... \circ \Phi_{G_{n-1}}^{\lambda^{n-1}} \circ \Phi_{G_n}^{\lambda^n} \quad . \tag{A.8}$$

The n Hamiltonians $G_1, ..., G_n$ are the so-called *generating Hamiltonians* of the canonical transformation $\mathcal{C}_\lambda$; said transformation is completely specified when all the generating Hamiltonians are known. In fact, for the perturbative approach introduced here, the generating Hamiltonian are precisely the unknowns of the process, and their form is determined order by order. Actually, we will see that the normal form of a given quasi-integrable Hamiltonian H_λ is not unique, but there are infinitely many possible sets of generating Hamiltonians $G_1, ..., G_n$ that bring the perturbation to normal form.

Let us now introduce some of the notation that we will make extensive use of. We start with the definitions of *time-average* and *deviation*:

Definition A.1.2. (Time-average) Given a Hamiltonian function h, the *time-average of f along the flow of h* is

$$\langle f \rangle_h := \frac{1}{t} \lim_{t\to\infty} \int_0^t (f \circ \Phi_h^s) ds \ , \tag{A.9}$$

where f is any real function defined on the phase space Γ .

Definition A.1.3. (Deviation) Given a Hamiltonian function h, the *deviation of f from its time-average* is

$$\delta_h f := f - \langle f \rangle_h \ , \tag{A.10}$$

where f is any real function defined on the phase space Γ .

Let us also state — and prove — the following technical lemmas, which we will later need.

Lemma A.1.4. *The time-average along the flow of the Hamiltonian function h is invariant with respect to the flow of h , i.e. for any function f one has*

$$\langle f\rangle_h \circ \Phi_h^t = \langle f\rangle_h \quad \forall t \ . \tag{A.11}$$

Proof. According to the definition, and making use of the group property of the flow $\Phi_h^s \circ \Phi_h^r = \Phi_h^{s+r}$, one obtains

$$\langle f\rangle_h \circ \Phi_h^r := \frac{1}{t} \lim_{t\to\infty} \int_0^t (f \circ \Phi_h^{s+r})\ ds = \lim_{t\to\infty} \frac{1}{t} \int_r^{t+r} f \circ \Phi_h^u\ du \ .$$

Let us split the integral in three parts, i.e. $\int_r^{t+r} du = \int_r^0 du + \int_0^t du + \int_t^{t+r} du$: the first and the third integrals are on bounded intervals (with lenght equal to r) , so the only term which survives is the second one. As a result one has

$$\langle f\rangle_h \circ \Phi_h^r(\xi) = \lim_{t\to\infty} \frac{1}{t} \int_0^t f(\Phi_h^u(\xi))\ du = \langle f\rangle_h \ .$$

□

Lemma A.1.5. *Let G be a Hamiltonian function, with Φ_G^s its flow. Then, for any function f,*

$$f \circ \Phi_G^s = e^{sL_g} f \ , \tag{A.12}$$

where L_G is the Lie derivative along the Hamiltonian vector field X_G, defined by

$$L_G := \{\cdot, G\} = (J\nabla G)\cdot\nabla = X_G \cdot \nabla \ . \tag{A.13}$$

Proof. Set $\widetilde{f}(s) := f \circ \Phi_G^s$ and notice that $\widetilde{f}(0) = f$.
One has

$$\dot{\widetilde{f}} = \{f, G\} \circ \Phi_G^s = \widetilde{L_G f} \ ,$$

so that $\ddot{\widetilde{f}} = \widetilde{L_G^2 f}$ and so on, i.e. $\frac{d^n \widetilde{f}}{ds^n} = \widetilde{L_G^n f}$ for any $n \geq 0$. Performing now the Taylor expansion of $\widetilde{f}$ at $s = 0$ one finally gets

$$\widetilde{f}(s) = \sum_{n\geq 0} \frac{s^n}{n!} \frac{d^n \widetilde{f}}{ds^n}|_{s=0} = \sum_{n\geq 0} \frac{s^n L_G^n}{n!} f = e^{sL_G f} \ .$$

□

Lemma A.1.6. *For any function f , the solution of the equation*

$$L_h g = \delta_h f \iff \{g, h\} = f - \langle f \rangle_h \tag{A.14}$$

is given by

$$g = \mathcal{G} + L_h^{-1}\delta_h f := \mathcal{G} + \lim_{t\to\infty} \frac{1}{t}\int_0^t (s-t)(\delta_h f \circ \Phi_h^s)\ ds\ , \tag{A.15}$$

where $\mathcal{G} \in ker L_h$, i.e. $\{\mathcal{G}, h\} = 0$.

Proof. Consider first the left part of (A.14) . Composing now both sides with the flow of h , multiplying by $(s-t)$, integrating with respect to s from 0 to t and dividing all this by t one obtains

$$\frac{1}{t}\int_0^t (s-t)(L_h g \circ \Phi_h^s)\ ds = \frac{1}{t}\int_0^t (s-t)(\delta_h f \circ \Phi_h^s)\ ds\ .$$

One notices that Lemma (A.1.5) leads to

$$g = \frac{1}{t}\int_0^t e^{sL_h} g\ ds + \frac{1}{t}(s-t)\delta_h f \circ \Phi_h^s\ ds\ ;$$

performing the limit for $t \to \infty$ of the latter equation and observing that $\mathcal{G} := \langle g \rangle_h \in \ker L_h$, thanks to Lemma (A.1.4) the thesis follows. □

Theorem A.1.7. ***(Averaging Principle)*** *Consider a quasi-integrable Hamiltonian H_λ of the form* (A.1) *satisfying the three hypothesis made above; then*

1. *For any choice of the generating Hamiltonians $G_1, ..., G_n$, defining the canonical transformation* (A.8) *, one has*

$$\tilde{H_\lambda} = H_\lambda \circ \mathcal{C}_\lambda^{-1} = h + \sum_{j=1}^{n} \lambda^j \mathcal{P}_j + \mathcal{R}_{n+1}\ , \tag{A.16}$$

   where, for every $j = 1, ..., n$ and taking $F_1 = 0$

$$\mathcal{P}_j = -L_h G_j + P_j + F_j[h, P_1, ..., P_{j-1}, G_1, ..., G_{j-1}]\ , \tag{A.17}$$

$$\mathcal{R}_{n+1} = \lambda^{n+1}(P_{n+1} + F_{n+1}[h, P_1, ..., P_n, G_1, ..., G_n]) + O(\lambda^{n+2})\ . \tag{A.18}$$

2. *The perturbation at order $j = 1, ..., n$ of the normal form is given by*

$$S_j = \langle P_j + F_j \rangle_h\ . \tag{A.19}$$

3. The j^{th} generating Hamiltonian, for $j = 1, ..., n$, is given by

$$G_j = \mathcal{G}_j + L_h^{-1}\delta_h(P_j + F_j) \tag{A.20}$$

where $\mathcal{G}_j \in ker L_h$, $\forall j$.

Proof. Let us denote $L_j = L_{G_j} = \{\cdot, G_j\}$ $\forall j = 1, ..., n$. According to Lemma (A.1.5) and to the definition of the canonical transformation $\mathcal{C}_\lambda^{-1}$ one has

$$\tilde{H}_\lambda = e^{\lambda^n L_n} \cdots e^{\lambda^2 L_2} e^{\lambda L_1}(h + \lambda P_1 + \lambda^2 P_2 + ... + \lambda^n P_n + R_{n+1}) \ . \tag{A.21}$$

We are going to prove the statements by induction, first considering the transformation for the case $n = 2$.
Performing the expansion of the exponentials what comes out is

$$\begin{aligned}\tilde{H}_\lambda = h + \lambda(L_1 h + P_1) + \lambda^2(L_2 h + P_2 + L_1 P_1 + \frac{1}{2}L_1^2 h) + \\ + \lambda^3(P_3 + L_2 P_1 + \frac{1}{2}L_1^2 P_2 + \frac{1}{6}L_1^3 h + L_2 L_1 h) + O(\lambda^4) \ ,\end{aligned}$$

which satisfies (A.17)-(A.18) , i.e.

$$\begin{aligned}&\mathcal{P}_1 = -L_h G_1 + P_1 \ , \quad F_1 = 0 \ ; \\ &\mathcal{P}_2 = -L_h G_2 + P_2 + F_2 \ , \quad F_2 = L_1 P_1 + \frac{1}{2}L_1^2 h \ , \\ &\mathcal{R}_3 = \lambda^3(P_3 + F_3) + O(\lambda^4) \ , \quad F_3 = L_2 P_1 + \frac{1}{2}L_1^2 P_2 + \frac{1}{6}L_1^3 h + L_2 L_1 h \ .\end{aligned}$$

Statement (1) : it holds for $n = 1, 2$. Let us denote with $\tilde{H}_\lambda^{(m-1)}$ the Hamiltonian up to the order $m-1$ and suppose that it has the form (A.17)-(A.18). Take the exponential operator $e^{\lambda^m L_m} = 1 + \lambda^m L_m + O(\lambda^{2m})$ and let it act on the left of $\tilde{H}_\lambda^{(m-1)}$, obtaining $\tilde{H}_\lambda^{(m)} = e^{\lambda^m L_m}\tilde{H}_\lambda^{(m-1)}$. At this point one has

$$\tilde{H}_\lambda^{(m)} = h + \sum_{j=1}^{m} \lambda^j \mathcal{P}_j + \mathcal{R}_{m+1}$$

with

$$\mathcal{R}_{m+1} = \lambda^m(-L_h G_m + P_m + F_m) + O(\lambda^{m+1}) \ ,$$

which has the form we were looking for and by induction it holds for any j up to any pre-fixed order n.

Statement (2 - 3) : in order to get the Hamiltonian (A.16) in normal form we impose the condition $\mathcal{P}_j = S_j \in \ker L_h$, where $\mathcal{P}_j$ is given by (A.17) .

At this point one has to solve the **homological equation** of Hamiltonian perturbation theory

$$S_j = -L_h G_j + P_j + F_j \tag{A.22}$$

with respect to the unknows S_j and G_j .
Composing both sides of (A.22) with $\Phi_h^{s\ 1}$, integrating with respect to s from 0 to t and dividing all by t one gets

$$S_j = -\frac{G_j \circ \Phi_h^t - G_j}{t} + \frac{1}{t}\int_0^t (P_j + F_j) \circ \Phi_h^s \, ds \ ,$$

so statement (2) follows in the limit $t \to \infty$ by the boundedness of the unperturbed flow.
Now put into (A.22) the expression $S_j = \langle P_j + F_j \rangle_h$, which yields

$$L_h G_j = P_j + F_j - \langle P_j + F_j \rangle_h = \delta_h (P_j + F_j) \ ,$$

so statement (3) holds thanks to Lemma (A.1.6) , which reveals that the solution of the latter equation is

$$G_j = \mathcal{G}_j + L_h^{-1} \delta_h (P_j + F_j) \ .$$

□

A.2 Quantum perturbation theory

The perturbation theory formalism in the work frame of classical mechanics can be automatically translated in a quantum environment. This is possible thanks to the Poisson structure of the algebra of linear self-adjoint operators in the Heisenberg picture, i.e. the existence of a proper Poisson bracket acting on the operators (the commutator over $i\hbar$). The formalism of the Hamiltonian perturbation theory can be applied to quantum mechanics regardless of the choice of coordinates, just like in the classical case. To see why, we first notice that unitary transformations of the wave function ψ — the unknown of the Schrödinger equation — , act as canonical transformations; this is true, since the Schrödinger equation performs the role of the classical Hamilton equations, and is preserved precisely by unitary transformations. Explicitly, let us take the map

$$\psi \longmapsto \psi' = \hat{U}^\dagger \psi \ , \quad \hat{H} \longmapsto \hat{H}' = \hat{U}^\dagger \hat{H} \hat{U} \quad , \quad \hat{U}^\dagger \hat{U} = \mathbb{1} \ , \tag{A.23}$$

[1] In the composition we take into account that $S_j \circ \Phi_h^s = S_j$ and $L_h G_j \circ \Phi_h^s = \frac{d(G_j \circ \Phi_h^s)}{ds}$.

so that the Schrödinger equation

$$i\hbar\psi_t = \hat{H}\psi \tag{A.24}$$

becomes

$$i\hbar\psi'_t = \hat{H}'\psi' \ , \tag{A.25}$$

which is formally the same, as anticipated.

Knowing this, we can now look at the perturbative context, in which the Hamiltonian is of the form $\hat{H} = \hat{h} + \lambda\hat{P}$. Naturally, the Schrödinger equation will then be

$$i\hbar\psi_t = \hat{H}_\lambda\psi = (\hat{h} + \lambda\hat{P})\psi \quad , \tag{A.26}$$

where $\hat{H}_\lambda$ is a certain self-adjoint operator consisting of an unperturbed part $\hat{h}$ and a perturbation $\hat{P}$, with λ a small parameter. Just like in the classical case, one can try to eliminate the perturbation $\hat{P}$ by looking for a particular unitary time-dependent operator $\hat{U}(\lambda)$ that conjugates $\hat{H}$ to its normal form. Therefore, we work once again by analogy with the classical case, imagining the canonical transformation $\hat{U}(\lambda)$ to be the Schrödinger flow (i.e. a time evolution operator) at time λ of some unknown Hamiltonian, the generator, which is going to be a Hermitian operator $\hat{G}$. Namely, the transformation will be

$$\hat{U}_\lambda = e^{\frac{1}{i\hbar}\lambda\hat{G}} \ . \tag{A.27}$$

To any Hermitian operator $\hat{G}$ one can associate an operator $\mathcal{L}_{\hat{G}}$, which plays the role of the quantum Lie derivative along $\hat{G}$, defined as the following:

$$\mathcal{L}_{\hat{G}} := \frac{1}{i\hbar}[\ \cdot\ , \hat{G}] = [\![\ \cdot\ , \hat{G}]\!] \ , \tag{A.28}$$

Furthermore, it is easy to obtain the quantum version of Lemma (A.1.5), which is the following

Lemma A.2.1. *Let $\hat{F}$ and $\hat{G}$ be two Hermitian λ-independent operators. Then*

$$e^{i\lambda\hat{G}/\hbar}\hat{F}e^{-i\lambda\hat{G}/\hbar} = e^{\lambda\mathcal{L}_{\hat{G}}}\hat{F} \ . \tag{A.29}$$

Proof. Let $\hat{F}(\lambda)$ define the left hand side of the equation, and take its derivative with respect to λ:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\hat{F}(\lambda) = \frac{i}{\hbar}\big(\hat{G}\hat{F}(\lambda) - \hat{F}(\lambda)\hat{G}\big) = \mathcal{L}_{\hat{G}}\hat{F}(\lambda) \quad .$$

This equation can then be formally integrated with initial condition $\hat{F}(0) = \hat{F}$, giving

$$\hat{F}(\lambda) = e^{\lambda\mathcal{L}_{\hat{G}}}\hat{F} \ .$$

□

From the Lemma, it follows that the operator $\hat{F}$ is invariant with respect to the flow of $\hat{G}$ if and only if $[\hat{F}, \hat{G}] = 0$. What we proved here is simply the Heisenberg evolution equation for observables, i.e. for a generic operator $\hat{A}$,

$$\frac{d}{dt}\hat{A}_{\hat{H}}(t) = [\![\hat{A}_{\hat{H}}(t)\ ,\ \hat{H}]\!] = \frac{1}{i\hbar}[\hat{A}_{\hat{H}}(t)\ ,\ \hat{H}] = \mathcal{L}_{\hat{H}}\hat{A}_{\hat{H}}(t)\ .$$

We have shown that, starting from the Schrödinger equation, one can find an equivalence to the classical Hamiltonian equation for evolution by shifting the focus onto the quantum observables, the operators, with the Heisenberg picture evolution. This parallel is nice and complete, as the quantum Heisenberg Lie derivative (A.28) picks up the role of proper Poisson brackets, giving birth to a well defined Poisson structure for the system. Thanks to this, we are able to exactly retrace the steps leading to the classical Averaging Principle, and to quickly find the quantum version of the normal form construction.

A.2.1 The quantum normal form

Let us again think of a quantum Hamiltonian in the following form:

$$\hat{H}_\lambda = \hat{h} + \sum_{j=1}^{n} \lambda^j \hat{P}_j \quad . \tag{A.30}$$

We are interested on its normal form up to a fixed order. Let us very quickly see how the classical procedure for finding the normal form is kept conceptually the same up to first order; higher orders are left to the reader, but it will be evident that there is really nothing different to do with respect to the classical approach.

As we have seen above, the proper change of coordinate can be achieved by looking for a unitary transformation $\hat{U}_{\hat{G}_1}(\lambda)$ generating by an unknown Hamiltonian (Hermitian) operator $\hat{G}_1$ at time λ. Therefore, we define $\hat{U}_1 := \hat{U}_{\hat{G}_1}$, which will look like

$$\hat{U}_1^\dagger(\lambda) = e^{-\frac{1}{i\hbar}\lambda\hat{G}_1} \quad ; \tag{A.31}$$

we require that, under the action of $\hat{U}_1$, the Hamiltonian (A.30) is sent to

$$\begin{aligned} \hat{\mathcal{H}}_\lambda &= e^{\lambda\mathcal{L}_1}\hat{H}_\lambda = \hat{h} + \lambda\hat{S}_1 + ... \\ &= \hat{h} + \lambda\big(\hat{P}_1 + \mathcal{L}_{\hat{G}}\hat{h}\big) + \mathcal{O}(\lambda^2) \quad , \end{aligned} \tag{A.32}$$

where $\mathcal{L}_1 = \mathcal{L}_{\hat{G}_1} = \frac{1}{i\hbar}[\cdot, \hat{G}_1]$ and $\hat{S}_1$ is such that $[\hat{S}_1, \hat{h}] = 0$. Notice the formal equivalence with the classical version. Taking into account that $\mathcal{L}_{\hat{G}}\hat{h} =$

$-\mathcal{L}_{\hat{h}}\hat{G}$, we can write down the first order quantum homological equation

$$\hat{S}_1 = P_1 - \mathcal{L}_{\hat{h}}\hat{G}_1 \tag{A.33}$$

and solve it for $\hat{S}_1$ and $\hat{G}_1$, just like in the classical case. The resulting solutions are

$$\hat{S}_1 = \langle \hat{P}_1 \rangle_{\hat{h}} = \lim_{t\to\infty} \frac{1}{t} \int_0^t \hat{U}^\dagger_{\hat{h}}(s)\hat{P}_1\hat{U}_{\hat{h}}(s)\ ds \tag{A.34}$$

and

$$\hat{G}_1 = \hat{\mathcal{G}}_1 + \lim_{t\to\infty} \frac{1}{t} \int_0^t (s-t)\hat{U}^\dagger_{\hat{h}}(s)(\hat{P}_1 - \langle \hat{P}_1 \rangle_{\hat{h}})\hat{U}_{\hat{h}}(s)\ ds \quad , \tag{A.35}$$

where $\hat{\mathcal{G}}_1$ is any Hermitian operator such that $[\hat{\mathcal{G}}_1, \hat{h}] = 0$. The formulas for the next orders can be obtained by using the quantum Averaging Principle. We simply report here the ones for the second order,

$$\hat{S}_2 = \langle \hat{P}_2 + \hat{F}_2 \rangle_{\hat{h}} \ , \tag{A.36}$$

$$\hat{F}_2 = \mathcal{L}_1\hat{P}_1 + \frac{1}{2}\mathcal{L}_1^2\hat{P}_1 \ , \tag{A.37}$$

and

$$\hat{G}_2 = \hat{\mathcal{G}}_2 + \mathcal{L}^{-1}_{\hat{h}}\delta_h(\hat{P}_2 + \hat{F}_2) \ , \tag{A.38}$$

where $[\mathcal{G}_2, \hat{h}] = 0$.

We have seen, then, that the Hamiltonian perturbation theory represents a powerful tool that can help study a great deal of systems which are not integrable, as long as they remain similar to other integrable ones, in the sense described before. It is widely used to study classical systems, but it may become a precious approach to understand quantum systems as well, since there is no conceptual difference between the classical and quantum case. Moreover, as opposed to the usual perturbation theory in quantum mechanics, which has a strictly static point of view, this Hamiltonian approach allows us to keep a watch over the dynamic of the system, a dynamic which is itself a pertubed one.

Appendix B

The Bethe Ansatz

In 1931, Hans Bethe published in [11] a method for solving the one dimensional spin-1/2 Heisenberg model by obtaining the exact eigenvalues. It consists in a parametrization of the eigenvectors, the so called *Bethe ansatz*, which, beyond Bethe's expectations, was found to be actually expendable in the case of other one-dimensional systems. Indeed, many other quantum many body systems are known to be solvable by some variant of the Bethe ansatz.

Since the Bethe ansatz is rarely discussed in textbooks, except at the advanced level, we add this appendix in order to introduce the Bethe ansatz at an elementary level, following [13], [12], [14]. Although the eigenvalues and eigenvectors for a finite system may be more easily obtained by numerical diagonalization, the Bethe ansatz bears some important advantages. Among them, there is the big asset of being able, in many cases, to evaluate the eigenvalues and the physical properties in the thermodynamic limit. Indeed, the Bethe ansatz is a basis transformation that does not have to be supported by a numerical diagonalization, in a way that removes the cap on system size. However, as we will see, the approach comes with some significant computational and analytical challenges.

B.1 Vacuum state and magnon excitations

Let us start with the study of the vacuum state and the first excitation.

The $M = 0$ case is the case in which all spins are aligned; let us say that this vacuum state is the state where all spins are down. This is just a convention, many sources regard their vacuum state as the one with all spins up. So, the vacuum state, the single vector $|0\rangle$, is an eigenstate, $H|0\rangle =$

$E_0 |0\rangle$, with energy

$$|0\rangle = |\downarrow ... \downarrow\rangle \quad \Longrightarrow \quad E_0 = \frac{-g}{4} N \quad . \tag{B.1}$$

The natural way to describe the invariant subspace in which $M = 1$ (one spin down) is to label the position of the flipped spin:

$$|n\rangle = S_n^+ |0\rangle \ , \quad n = 1, ..., N \quad . \tag{B.2}$$

Such states are clearly not eigenstates of H because of the coupling term, but we can construct N new states as linear combinations of $|n\rangle$, simply

$$|\psi(k)\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} a_n(k) |n\rangle \quad , \tag{B.3}$$

for wave numbers $k = 2\pi m/N$,$m = 0, ..., N-1$. The ansatz given by Bethe in this case is a guess for the value of the coefficients $a_n(k)$, keeping in mind that they should respect translational symmetry, i.e. the invariance of H with respect to discrete translations. The Bethe states are, explicitly,

$$|\psi(k)\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} e^{ikn} |n\rangle \quad . \tag{B.4}$$

The vectors $|\psi\rangle$ are, as we requested, eigenstates of the translation operator with eigenvalues e^{ik}; their crucial property is that they are also eigenstates of H with eigenvalues

$$E(k) = E_0 + g(1 - \cos k) \quad , \tag{B.5}$$

as can be verified. The vectors (B.4) represent the so called *magnon excitations* ($\Delta S = 1$ excitations), which means that the complete spin down alignment of the polarized vacuum state $|0\rangle$ is periodically disturbed by a single spin wave with wavelength $\lambda = 2\pi/k$ (the magnon). Note that the $k = 0$ state, or the $k = 2\pi$ if $m = 1, ..., N$, is degenerate with $|0\rangle$.

B.2 The two-body problem

At first, one might think the $M > 1$ case to be just a simple superposition of magnons. Alas, just by comparing the number of the states in each base, we see that this is not the case. Let us discuss in detail the $M = 2$ case.

Again, we can think of a new state as a generic linear combination of the two-spin-flip base states $|n_1, n_2\rangle = S^+_{n_1} S^+_{n_2} |0\rangle$, i.e.

$$\psi(k_1, k_2) = \sum_{1 \le n_1 < n_2 \le N} a_{n_1,n_2}(k_1, k_2) \, |n_1, n_2\rangle \quad . \tag{B.6}$$

With this, the eigenvalue equation translates into:

$$2[E - E_0] a_{n_1,n_2} = g[4a_{n_1,n_2} - a_{n_1-1,n_2} - a_{n_1+1,n_2} - a_{n_1,n_2-1} - a_{n_1,n_2+1}] \quad , \tag{B.7}$$
$$(\text{for } n_2 > n_1 + 1)$$

$$2[E - E_0] a_{n_1,n_2} = g[2a_{n_1,n_2} - a_{n_1-1,n_2} - a_{n_1,n_2+1}] \quad , \tag{B.8}$$
$$(\text{for } n_2 = n_1 + 1)$$

Bethe's preliminary ansatz to determine the coefficients $a_{n_1,n_2}(k_1, k_2)$ has been

$$a_{n_1,n_2}(k_1, k_2) = A e^{i(k_1 n_1 + k_2 n_2)} + A' e^{i(k_1 n_2 + k_2 n_1)} \quad , \tag{B.9}$$

which, once again, automatically satisfies the eigenvalue equations with energy

$$E(k_1, k_2) = E_0 + g \sum_{j=1,2} (1 - \cos k_j) \quad . \tag{B.10}$$

From the eigenvalue equation we can also recover the scattering phase relation for the coefficients A and A' of the ansatz:

$$\frac{A}{A'} := e^{i\theta} = \frac{e^{i(k_1+k_2)} + 1 - 2e^{ik_1}}{e^{i(k_1+k_2)} + 1 - 2e^{ik_2}} \quad , \tag{B.11}$$

which can also be restated in the form

$$2 \cot \frac{\theta}{2} = \cot \frac{k_1}{2} - \cot \frac{k_2}{2} \quad . \tag{B.12}$$

Now, the quasi-momenta k_1, k_2 of the Bethe Ansatz wave function are not simply the magnon wave numbers; they are related to those by the phase θ and can be determined by requiring that the wave function $\psi(k_1, k_2)$ satisfies the periodic boundary conditions $a_{n1,n2} = a_{n2,n1+N}$, i.e.

$$e^{ik_1 N} = e^{i\theta} \, , \quad e^{ik_2 N} = e^{-i\theta} \, . \tag{B.13}$$

Equivalently,

$$k_1 = \frac{2\pi m_1}{N} + \frac{\theta}{N} \ , \quad k_2 = \frac{2\pi m_2}{N} - \frac{\theta}{N} \ . \tag{B.14}$$

where the $m_j = 0, ..., N-1$ are the usual integer quantum numbers.

The total momentum of this state is independent of θ and reads

$$K = k_1 + k_2 = \frac{2\pi(m_1 + m_2)}{N} \quad . \tag{B.15}$$

The fact that the magnons in this state are interacting is reflected in the phase shift θ, i.e. in the deviation of the quasi-momenta $k_{1,2}$ from the free magnon wave numbers. It has to be said that the quasi momenta k_1, k_2 are necessary to specify the Bethe Ansatz wave function but are not observable (they may be regarded as bookkeeping parameters), while the wave number K is the quantum number associated with the translational symmetry of H and exists independently of the Bethe Ansatz.

The allowed (m_1, m_2) pairs are restricted to $0 \leq m_1 \leq m_2 \leq N-1$, because switching the two simply interchanges k_1 and k_2 and produces virtually the same solution. There are $N(N+1)/2$ pairs that meet the ordering restriction, but only $N(N+1)/2$ of them are compatible as a solution for equation (B.14). Note that the scattering phase θ does not depend on the difference between the momenta of the scattering magnons. This means that equal Bethe numbers $m_1 = m_2$ do not imply $k_1 = k_2$ (which would make the coefficients of the wavefunction ψ vanish). Thus, we cannot exclude solutions with equal quantum numbers, and therefore we lack of a good criterion to exclude the N spurious choices of Bethe numbers.

The solutions for the interacting Bethe wave numbers can be determined analytically or computationally. Some of them have real k_1, k_2, while others yield complex conjugate momenta, $k_2 = k_1^*$. The majority of solutions are real and different from zero. Traditionally, they are classified in three separate cases:

- *C_1-class of states* — If $m_1 = 0$, then all solutions are real and $k_1 = 0$, $k_2 = 2\pi m_2/N$ for $m_2 = 0, ..., N-1$, $\theta = 0$. These states are clearly degenerate with the single magnon states.

- *C_2-class of states* — both m_1, m_2 are non-zero and are set apart by two or more, i.e. $m_2 - m_1 \geq 2$. There are $N(N-5)/2 + 3$ such pairs of numbers, and all of them are real and correspond to suitable quasi momenta k_1, k_2. To find them, we tweak equation (B.14) into a nonlinear equation for k_1, in terms of the total momentum K:

$$2 \cot \frac{Nk_1}{2} = \cot \frac{k_1}{2} - \cot \frac{K - k_1}{2} \quad . \tag{B.16}$$

Considering that the total momentum of the state is quantized ($K = 2\pi n/N$), we can substitute for different n to determine k_1 and $k_2 = K - k1$. There are several cases where the analytic solution is possible, while the others are usually found numerically.

- *C_3-class of states* — both m_1, m_2 are non-zero and are either equal or differ by 1. There are $2N - 3$ pairs possible, but only $N - 3$ of them are compatible with (B.14), and the solutions, when present, are complex. It is quite difficult to numerically find the complex solutions of a system of equations, so we are better off turning everything into a real equation by parameterizing in the following way:

$$k_1 := \frac{K}{2} + ik \ , \quad k_2 := \frac{K}{2} - ik \ . \tag{B.17}$$

This yields, for θ,

$$\theta = \pi(m_2 - m_1) + iNk \quad , \tag{B.18}$$

and

$$\cos\frac{K}{2}\sinh(Nk) = \sinh[(N-1)k] + \cos[\pi(m_1 - m_2)]\sinh k \quad . \tag{B.19}$$

which gives k as a function of the total momentum $K = 2\pi(m_1+m_2)/N$. This solution represents a bound state in which the two flipped spins cannot be more than order of k sites apart. Substituting the parametric equations into the energy formula we find the energy of this complex solution to be

$$E = E_0 + 2g\big(1 - \cos\frac{K}{2}\cosh k\big) \quad . \tag{B.20}$$

For $N \to \infty$, we get

$$k_{1,2} = \frac{K}{2} \pm i \ln\cos\frac{K}{2} \quad , \tag{B.21}$$

which means that for large systems the energy of the bound state is

$$E(N \to \infty) = E_0 + \frac{g}{2}(1 - \cos K) \quad . \tag{B.22}$$

In the large N limit, the real solutions for the k's are not too different from a simple superpositions of two magnons, as the quasi-momentum of each excitation differs from the free one as

$$k_{1,2} = \frac{2\pi}{N}m_{1,2} + \mathcal{O}\Big(\frac{1}{N^2}\Big) \quad . \tag{B.23}$$

B.3 The Bethe Solution

In analogy with the method used for the two-body problem, we can now build the eigenstates in the case of a generic number M of overturned spins. Just like we did in the previous section, we write the generic states as a linear combination of the states in the natural basis for fixed M:

$$|\psi(k_1,...,k_M)\rangle = \sum_{1\leq n_1<...<n_M\leq N} a_{n_1,...,n_M}(k_1,...,k_M)\,|n_1,...,n_M\rangle \quad . \tag{B.24}$$

Trying to generalize the form of the coefficients for a generic M case, starting from their form and the known solutions in the two-body case, we write them as such:

$$a_{n_1,...,n_M}(k_1,...,k_M) = \sum_{\sigma} \exp\Big[i\sum_{j=1}^{M} k_{\sigma j} n_j + \frac{i}{2}\sum l < j\theta(k_{\sigma l}, k_{\sigma j}) \quad , \tag{B.25}$$

where the sum runs over all $M!$ permutations σ of the assignments of the quasi-momenta k_j to each overturned spin n_j. We also introduced an antisymmetric phase shift, $\theta(k_l, k_j) = -\theta(k_j, k_l)$, which plays the role of a generalization of the θ phase shift in the two excitations case. The consistency equations for the coefficients $a_{n_1,...,n_M}(k_1,...,k_M)$ are once again found plugging the state into the eigenvalue equation $H\,|\psi\rangle = E\,|\psi\rangle$. They are, naturally, just a straightforward generalization of the two-particle case. The energy eigenvalue equation becomes

$$E(k_1,...,k_M) = E_0 + g\sum_{j=1}^{M}(1-\cos k_j) \quad , \tag{B.26}$$

and the eigenstate condition can be written as

$$2a_{n_1,...,n_j,n_{j+1},...,n_M} = a_{n_1,...,n_j,n_j,...,n_M} + a_{n_1,...,n_{j+1},n_{j+1},...,n_M} \quad , \tag{B.27}$$

for $j = 1,...,M$. These conditions fix the phase shift $\theta(k_j, k_l)$, similarly to the old one:

$$e^{i\theta(k_j,k_l)} = -\frac{e^{i(k_j+k_l)} + 1 - 2e^{ik_j}}{e^{i(k_j+k_l)} + 1 - 2e^{ik_l}} \quad . \tag{B.28}$$

Just like in the $M = 2$ case, this last set of equations can be presented in real form, once again in terms of cotangents:

$$2\cot\frac{\theta(k_j,k_l)}{2} = \cot\frac{k_j}{2} - \cot\frac{k_l}{2} \quad , \quad j,l = 1,...,M\ . \tag{B.29}$$

The periodic conditions on the chain for translations by N sites $a_{n_1,...,n_M} = a_{n_2,...,n_M,n_1+N}$ pose the constraint

$$\sum_{j=1}^{M} k_{\sigma_j} n_j + \frac{1}{2}\sum_{l<j} \theta(k_{\sigma_l}, k_{\sigma_j}) = \frac{1}{2}\sum_{l<j} \theta(k_{\sigma'_l}, k_{\sigma'_j}) - 2\pi m_{\sigma'_M} + \sum_{j=2}^{M} k_{\sigma'_{j-1}} n_j + k_{\sigma'_M}(n_1+N) \quad . \tag{B.30}$$

Here, the permutations σ' are defined as

$$\begin{cases} \sigma'(j-1) = \sigma(j) \ , \ j = 2, ..., M \\ \sigma'(M) = \sigma(1) \ . \end{cases} \tag{B.31}$$

The numbers m_j are integers and play the same role of m_1, m_2 in the previous section. From (B.31), it is clear that all terms which do not involve $\sigma'(M) = \sigma(1)$ are equal on both sides, and they cancel. Thus, we get the simpler relations between the phase shifts and the quasi-momenta:

$$k_j = \frac{2\pi m_j}{N} + \frac{1}{N}\sum_{l\neq j} \theta(k_j, k_l) \quad , \quad j = 1, ..., M \ , \tag{B.32}$$

where $m_j \in \{0, 1, ..., N-1\}$, all in complete analogy with the two-body system.

Recall that even in the case of two flipped spins we had more to dwell upon, as not all choices of quantum numbers m_1, m_2 were suitable to produce solutions. Of course the generalization to m_j for $j = 1, ..., M$ is no different. Furthermore, among the possible choices of quantum numbers there are some that yield complex quasi-momenta, and this does nothing but increase the computational complexity of solving the Bethe equations. As a matter of fact, the classification of such complex solutions is so convoluted that it still presents an open problem. For more details on these issues, see [12]. A lot more can be said, though, in the case of thermodynamic limit $N \to \infty$, which simplifies the circumstances. In this case, we will see that it is possible to assume that all complex solutions organize themselves into strings — a so called *string hypothesis*. Such an assumption does provide an aid in constructing the Hilbert space of the system, and thus in studying its thermodynamic properties.

Introduction of rapidities

Before we delve into the string solutions, let us notice that the scattering phase $\theta(k_l, k_j)$ does not depend on the difference between the particle quasi-momenta. This means that we cannot exclude solutions with the same

quantum numbers, and more generally that we have more possible choices of quantum numbers than states in the Hilbert space. To fix this problem and restore translational invariance to the Bethe equations, we introduce the *rapidities* λ_j to parametrize the quasi-momenta in the following way:

$$\cot\frac{k_j}{2} = \lambda_j \ , \quad \text{i.e.} \quad k_j = \frac{1}{i}\ln\frac{\lambda_j + i}{\lambda_j - i} = \pi - \theta_1(\lambda_j) \quad , \tag{B.33}$$

where

$$\theta_n(\lambda_j) := 2\arctan\frac{\lambda_j}{n} \quad . \tag{B.34}$$

The energy and momentum of a single magnon associated to a quasi-momentum k are, in terms of the new variable λ,

$$p_0(\lambda) = k = \frac{1}{i}\ln\frac{\lambda + i}{\lambda - i} \quad , \tag{B.35}$$

$$\epsilon_0(\lambda) = g(1 - \cos k) = \frac{2g}{\lambda^2 + 1} \quad . \tag{B.36}$$

Here we have inserted a subscript 0 just to indicate that it corresponds to a single (real) particle — what we later be called a *0-type string*. The scattering phase becomes, again in terms of the rapidities,

$$\theta(k_j, k_l) = -\theta_2(\lambda_j - \lambda_l) + \pi\mathrm{sgn}[\Re(\lambda_j - \lambda_l)] \quad , \tag{B.37}$$

where $\Re(x)$ is the real part of x and $\mathrm{sgn}(y)$ means the sign of y. The Bethe equations for the k_j's set in terms of the rapidities are then

$$\theta_1(\lambda_j) = \frac{2\pi\tilde{m}_j}{N} + \frac{1}{N}\sum_{l=1}^{M}\theta_2(\lambda_j - \lambda_l) \quad , \quad j = 1, ..., M \ , \tag{B.38}$$

where the $\tilde{m}_j$ numbers are again integers, but different from before, and with different qualities. The state is now defined by these new Bethe numbers $\tilde{m}_j = 1, ..., M$, and while it is not easy to relate them to the previous quantum numbers, we will not need to do it and thus we will not be concerned with the matter. Because (B.38) is now translational invariant, two equal quantum numbers $\tilde{m}_j$ produce the same rapidities. Thus, we do not need to count sets of quantum numbers with repetitions. This is principally the reason why it is more convenient and widely accepted to work with the $\tilde{m}_j$, which now have "fermionic" properties, and thus provide the proper counting of the states (recall that each choice of increasing and non-repeating quantum

numbers produces a physical state). If we completely switch to rapidities to parametrize the eigenstates, we find that their energies and momenta are

$$E = E_0 + \sum_{j=1}^{M} \epsilon_0(\lambda_j) \quad , \tag{B.39}$$

$$K = \left[\sum_{j=1}^{M} p_0(\lambda_j) \right] \bmod 2\pi = \left[\pi M - \frac{2\pi}{N} \sum_{j=1}^{M} \tilde{m}_j \right] \bmod 2\pi \quad . \tag{B.40}$$

Let us recap: initially, we have expanded the wavefunction coefficients for the Bethe state as a sum of plane waves parametrized by quasi-momenta, but found that, as a result, the scattering phase was not explicitly translational invariant. The change of variable from quasi-momenta into rapidities shows that the basis $\left(\frac{\lambda+i}{\lambda-i}\right)^n$ is a more appropriate choice for the wavefunction ansatz, with as a consequence the knowledge that any complete single-particle basis can be used for the ansatz. Unfortunately, the problem of identifying and classifying real and complex solutions is not resolved by only turning to rapidities, as the computational complexity remains unchanged. However, we will see that the string hypothesis for the complex solutions can account for them in a very elegant way.

B.3.1 String solutions

Bethe himself had already noticed (in [11]) that the Bethe equations admit complex solutions, which in general are to be found numerically. Computationally, it might not be an easy task. However, there is a way around it if we take the thermodynamic limit $N \to \infty$. There is a simple structure that emerges in this limit; it is known as *string hypothesis*, [30]. The reason it is called a hypothesis is that it is not yet clear whether the resulting string solutions do in fact exhaust the whole Hilbert space. The situation is even more unsure for the case of XXZ chain. Despite the possible incompleteness of the solutions, there seems to be consensus about the fact that the string hypothesis does provide an accurate description of the thermodynamics of the system. This would suggest that solutions that are not part of the string structure, if they exist, might be relevant only for certain response functions or interesting only in the case of out-of-equilibrium studies [31].

Before we start with the general case, let us study again the case of two flipped spins $M = 2$.

The Bethe equations, written in terms of the rapidities, are:

$$\left(\frac{\lambda_1 + i}{\lambda_1 - i}\right)^N = \frac{\lambda_1 - \lambda_2 + 2i}{\lambda_1 - \lambda_2 - 2i} \quad , \tag{B.41}$$

$$\left(\frac{\lambda_2+i}{\lambda_2-i}\right)^N = \frac{\lambda_2-\lambda_1+2i}{\lambda_2-\lambda_1-2i} \quad . \tag{B.42}$$

If the imaginary part of λ_1, $\Im(\lambda_1)$, is different form zero, then the left hand side in (B.41) will vary exponentially with N. Therefore, in the thermodynamic limit it is either zero or infinity and the right hand side will have to follow, so either the numerator or the denominator must be zero. Thus, we must have

$$\lambda_1 - \lambda_2 = \pm 2i \ , \quad \text{i.e.} \quad \lambda_{1,2} = \lambda \pm i \ . \tag{B.43}$$

Then, the energy and momentum of this state are real:

$$p_{12}(\lambda) = p_0(\lambda+i) + p_0(\lambda-i) = \frac{1}{i}\ln\frac{\lambda+2i}{\lambda-2i} \quad , \tag{B.44}$$

$$\epsilon_{12}(\lambda) = \epsilon_0(\lambda+i) + \epsilon_0(\lambda-i) = \frac{4g}{\lambda^2+4} \quad , \tag{B.45}$$

which gives the dispersion relation

$$\epsilon_{12}(p) = \frac{g}{2}(1-\cos p_{12}) \quad . \tag{B.46}$$

Notice that for $g > 0$ we have

$$\epsilon_{12}(p) < \epsilon_0(p-p') + \epsilon_0(p') \quad , \tag{B.47}$$

meaning that in the ferromagnetic regime these bound states are energetically favored compared to real solutions.

For $M > 2$, as we said initially, we make the following assumption: complex solutions can be organized into so called *complexes* (or *strings*) of $2R+1$ rapidities, where all the rapidities have the same real value λ_R, and different, equidistant, imaginary parts. The new index runs by half integers, $R = 0, 1/2, 1, ...$, and the rapidities have the following form:

$$\lambda_r^{(R)} = \lambda_R + 2ir \ , \quad r = -R, -R+1, ..., R-1, R \ . \tag{B.48}$$

If we call ν_R the number of strings of length R, a state with a given magnetization satisfies

$$M = \sum_R (2R+1)\nu_R \quad . \tag{B.49}$$

We expect the results obtained through the string hypothesis to be a good approximation of reality as long as the number of single-particle solutions ν_0 dominates over all other parts in the sum.

These rapidities represent a group of $2R+1$ spins that move together with the same real rapidity. Also, and these spins are more likely to stay close to one-another — otherwise they would penalize the wavefunction by an exponential decay the farther away they are. Therefore, we will treat them as a single entity. Actually, all interactions between the individual rapidities in a string can be factorized and summed over separately when computing the interactions between the strings themselves. The energy and momentum of an R-complex are obtained by summing over all the rapidities in the complex:

$$p_R(\lambda_R) = \sum_{r=-R}^{R} p_0(\lambda_R - 2ir) = \frac{1}{i}\ln\frac{\lambda_R + i(2R+1)}{\lambda_R - i(2R+1)} = \pi - \theta_{2R+1}(\lambda_R) \quad , \tag{B.50}$$

$$\epsilon_R(\lambda_R) = \sum_{r=-R}^{R} \epsilon_0(\lambda_R - 2ir) = \frac{2g(2R+1)}{\lambda_R^2 + (2R+1)^2} = \frac{g}{2R+1}(1-\cos p_R) \quad . \tag{B.51}$$

Due to the equidistance between the imaginary parts of the rapidities, we are left with a highly regular structure which enables us to eliminate various terms. After doing so, we see that we can consider the scattering phase of an R-complex with a 0-complex, by taking the product with respect to all the particles in the given string. This way, we find

$$S_{0,R}(\lambda_0 - \lambda_R) = S_{0,R}(\lambda) = \frac{\lambda + i2R}{\lambda - i2R}\frac{\lambda + i2(R+1)}{\lambda - i2(R+1)} \quad . \tag{B.52}$$

Generalizing, the scattering of two strings of length R and R' is

$$S_{R,R'}(\lambda) = \prod_{L=|R-R'|}^{R+R'} S_{0,L}(\lambda) \quad . \tag{B.53}$$

Now that we set the notation straight, we wish to describe an eigenstate of the Heisenberg chain in terms of:

- the number of complexes ν_R for each R ,

- the rapidities of their center of mass $\lambda_{R,j}$, where $\lambda_{R,j}$ is the real part of the j-th string of lenght R, and $j = 1, ..., \nu_R$.

The Bethe equations for the strings are obtained by grouping all the rapidities $\lambda_j^{(R)}$ which are part of the same complex, to then compute the products

within each complex, in such a way that we are left only with the consistency conditions on their real centers $\lambda_{R,j}$, just so:

$$e^{ip_R(\lambda_{R,j})N} = \prod_{R'} \prod_{\substack{j' \\ (R',j')\neq(R,j)}}^{\nu_{R'}} S_{R,R'}(\lambda_{R,j} - \lambda_{R',j'}) \ , \quad \forall R \ , j = 1, ..., \nu_R \ . \quad \text{(B.54)}$$

Once again, we take the logarithm to find everything in terms of scattering phases. To account for the branches of the logarithms for each complex type, we introduce the half-integer quantum numbers $m_{R,j}$. Going back to the identity

$$\frac{1}{i} \ln \frac{\lambda + in}{\lambda - in} = \pi - 2 \arctan \frac{\lambda}{n} = \pi - \theta_n(\lambda) \quad , \quad \text{(B.55)}$$

we finally find

$$\theta_{2R+1}(\lambda_{R,j}) = \frac{2\pi}{N} m_{R,j} + \frac{1}{N} \sum_{(R',j')\neq(R,j)} \theta_{R,R'}(\lambda_{R,j} - \lambda_{R',j'}) \quad , \quad \text{(B.56)}$$

where

$$\theta_{R,R'}(\lambda) := \sum_{L=|R-R'|}^{R+R'} [\theta_{2L}(\lambda) + \theta_{2L+2}(\lambda)] \quad , \quad \text{(B.57)}$$

and the $L = 0$ is intended to be omitted. The equations (B.56) are called the *Bethe-Gaudin-Takahashi equations*.

In the string hypothesis, each state is then characterized by the number of strings ν_R and by the Bethe quantum numbers $m_{R,j}$ of each string type. Since the Hilbert space of a spin chain is limited, not all quantum numbers are allowed, as usual. The first thing to notice is that, within each complex, it must be that $m_{R,j} \neq m_{R,j'}$, in order to have a non-vanishing solution. Furthermore, since the momenta are constrained within a Brillouin zone (we are dealing with a lattice in real space), the fitting quantum numbers are bounded. Also, observe that a diverging rapidity $\lambda_R^{(\infty)} = \infty$ (which corresponds to a quasi-momentum at the edge of the Brillouin zone) has a fixed scattering phase with all other particles, because $\arctan(\pm\infty) \pm \pi/2$. Therefore, we find that $\lambda_R^{(\infty)}$ is given by the Bethe number

$$m_R^{(\infty)} = -\sum_{R'\neq R} [2\min(R, R') + 1]\nu'_R - \left(2R + \frac{1}{2}\right)(\nu_R - 1) + \frac{N}{2} \quad . \quad \text{(B.58)}$$

Mind that adding an R-complex shifts this boundary by $\frac{1}{2\pi}\theta_{R,R}(\infty) = 2R+\frac{1}{2}$, so the maximum quantum number which describes a finite rapidity, before it goes to the edges, is

$$m_R^{\max} = m_R^{(\infty)} - \left(2R + \frac{1}{2}\right) - \frac{1}{2} = \frac{N-1}{2} - \sum_{R'} J(R,R')\nu_{R'} \quad , \tag{B.59}$$

where

$$J(R,R') := \begin{cases} 2\min(R,R') + 1 \, , & R \neq R' \\ 2R + \frac{1}{2} \, , & R = R' \end{cases} \quad , \tag{B.60}$$

while the additional shift of $1/2$ in (B.59) is there to take into account that with each rapidity the Bethe numbers shift from integers to half-integers and vice-versa. Since all the scattering phases are odd functions of their argument, we have that

$$m_{R,\min} = -m_{R,\max} \quad , \tag{B.61}$$

which indicates that there are

$$P_R = 2m_R^{\max} + 1 = N - 2 - \sum_{R'} J(R,R')\nu_{R'} \tag{B.62}$$

vacancies for an R-complex. Clearly, the range of allowed values becomes narrower and narrower for complexes of any size if strings are added to the system. Using the results here, one can estimate the number of states accessible within the hypothesis implemented. One can show that such number scales like 2^N as we would like it to, [32], [33], since that means that only few states are possibly neglected in the string hypothesis framework. Such neglected states usually involve a large number of complex rapidities — typically a finite percentage of the number of sites N, which is here brought to infinity. Those rapidities are not organized in strings; even so, they are still somewhat able to satisfy the Bethe equations, because the exponential growth/decay previously mentioned on that left hand side is accidentally properly compensated on the right hand side [31]. As it was said already, these spurious states do not contribute significantly to the thermodynamics of the model. They are nonetheless important to determine the completeness of the Bethe solution and for other investigations, such as correlation functions, dynamical responses, or in working with finite systems.

A word on the ferromagnetic case

Up until now, we have studied the elementary excitations of the Heisenberg chain, providing a classification of its states. To continue on with the study of the properties of the energy states in the Bethe ansatz, one has to specify wether the regime is ferromagnetic or antiferromagnetic, since there is a dramatic difference between the two. Even though it is the most relevant for physical applications, we will not concern ourselves with the antiferromagnetic case here[1].

In the case of a ferromagnetic coupling, the state with all spins in the same direction $|0\rangle$ can be taken as the ground state. As we have seen from the Bethe energies, the ground state is degenerate with all the other members of the $S = N/2$ multiplet, i.e. states that are created from $|0\rangle$ by adding zero-momentum magnons. Right after, the lowest energy states are the individual magnons of momentum k along with bound states complexes, which have lower energy compared to superposition of magnon excitations — this is natural, since a ferromagnetic coupling promotes the clustering of overturned spins. The ground state can be then described as a magnon-vacuum, where there is a quadratic dispersion relation for the excitations.

[1]For the interested reader, see [13].

Bibliography

[1] R. Wieser. Quantum spin dynamics. *ArXiv e-prints 1410.6383*, October 2014.

[2] Ryoichi Kikuchi. On the minimum of magnetization reversal time. *Journal of Applied Physics*, 27(11):1352–1357, 1956.

[3] E. Brezin and J. Zinn-Justin, editors. *Field Theory Methods and Quantum Critical Phenomena*, Les Houches 1988, Proceedings, Fields, strings and critical phenomena. North-Holland, North-Holland, 1988.

[4] J. Liouville. Note sur l'intègration des èquations différentielles de la dynamique, prèsentèe au bureau des longitudes le 29 juin 1853. *Journal de Mathèmatiques Pures et Appliquèes*, pages 137–138, 1855.

[5] Susanne Reffert. Introduction to integrable models (lecture notes). ITP University of Bern, Fall 2011.

[6] L. Faddeev. *Instructive History of the Quantum Inverse Scattering Method*, pages 69–84. Springer Netherlands, Dordrecht, 1995.

[7] J. Hubbard. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 276(1365):238–257, 1963.

[8] Anjan Kundu. Quantum integrable systems: Construction, solution, algebraic aspect. *eprint arXiv:hep-th/9612046*, 1996.

[9] Giulio Bertoli. Quantum integrability classes (master thesis). Universiteit Van Amsterdam, August 2015.

[10] G. Müller. Quantum spin chains: Simple models with complex dynamics. *Zeitschrift fur Physik B Condensed Matter*, 68:149–159, June 1987.

[11] H. Bethe. Zur theorie der metalle. *Zeitschrift für Physik*, 71(3):205–226, Mar 1931.

[12] M. Karbach and G. Muller. Introduction to the Bethe ansatz I. *eprint arXiv:cond-mat/9809162*, September 1998.

[13] F. Franchini, editor. *An Introduction to Integrable Techniques for One-Dimensional Quantum Systems*, volume 940 of *Lecture Notes in Physics, Berlin Springer Verlag*, 2017.

[14] Vadim Ohanyan. Introduction to quantum spin chains (lecture notes). Yerevan State University.

[15] P. Jordan and E. Wigner. Über das paulische äquivalenzverbot. *Zeitschrift für Physik*, 47(9):631–651, Sep 1928.

[16] Daniel C. Mattis and Elliott H. Lieb. Exact solution of a many-fermion system and its associated boson field. *Journal of Mathematical Physics*, 6(2):304–312, 1965.

[17] G. Benfatto, G. Gallavotti, and V. Mastropietro. Renormalization group and the fermi surface in the luttinger model. *Phys. Rev. B*, 45:5468–5480, Mar 1992.

[18] A. Kitaev. Anyons in an exactly solved model and beyond. *Annals of Physics*, 321:2–111, January 2006.

[19] Y.-A. Chen, A. Kapustin, and D. Radičević. Exact bosonization in two spatial dimensions and a new class of lattice gauge theories. *Annals of Physics*, 393:234–253, June 2018.

[20] T. D. Schultz, D. C. Mattis, and E. H. Lieb. Two-dimensional ising model as a soluble problem of many fermions. *Rev. Mod. Phys.*, 36:856–871, Jul 1964.

[21] Antonio Ponno. Mathematical physics (lecture notes). Università degli Studi di Padova, June 2018.

[22] Askold M. Perelomov. *Integrable Systems of Classical Mechanics and Lie Algebras*. Number v. 1. Birkhäuser Basel, 1989.

[23] Michael A. Nielsen. The fermionic canonical commutation relations and the jordan-wigner transform (lecture notes). The University of Queensland, July 2005.

[24] L. Messio B. Doucot, B. Estienne. Problem: the jordan wigner transformation (tutorials). M2 ICFP, Theoretical Condensed Matter, 2017-2018.

[25] Elliott Lieb, Theodore Schultz, and Daniel Mattis. Two soluble models of an antiferromagnetic chain. *Annals of Physics*, 16(3):407 – 466, 1961.

[26] A.S. Davydov. *Theory of Solids*. Nauka, 1980.

[27] L. D. Faddeev and L. A. Takhtajan. *Hamiltonian Methods in the Theory of Solitons*. Springer Series in Soviet Mathematics, 1987.

[28] N.N. Bogolyubov jr. (originator). *Perturbation theory*. Encyclopedia of Mathematics.

[29] B. Eckhardt. Birkhoff-gustavson normal form in classical and quantum mechanics. *Journal of Physics A: Mathematical and General*, 19(15):2961, 1986.

[30] Minoru Takahashi and Masuo Suzuki. One-dimensional anisotropic heisenberg model at finite temperatures. *Progress of Theoretical Physics*, 48(6):2187–2209, 1972.

[31] Rob Hagemans and Jean-Sstien Caux. Deformed strings in the heisenberg model. *Journal of Physics A: Mathematical and Theoretical*, 40(49):14605, 2007.

[32] Minoru Takahashi. *Thermodynamics of One-Dimensional Solvable Models*. Cambridge University Press, 1999.

[33] L. D. Faddeev and L. A. Takhtadzhyan. *Spectrum and Scattering of Excitations in the the One-dimensional Isotropic Heisenberg Model*, pages 296–322.